MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

STATISTISCHE AFDELING

Leiding: Prof. Dr D. van Dantzig
Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

Rapport S 158

Elementaire statistische opgaven met uitgewerkte oplossingen

door

Prof.Dr J. Hemelrijk

en

Ir Doralien Wabeke

1954

INHOUD

	Blz.
Inleiding	1
Notaties	1
Formules	2
Opgaven	7
Oplossingen	22
Tabellen	73

INLEIDING

De bedoeling van deze oefeningenverzameling is tegemoet te komen aan de behoefte aan eenvoudig oefenmateriaal van elementair karakter. De opgenomen oefeningen zijn voor een groot deel ontleend aan tentamina van de Technische Hogeschool te Delft en sluiten aan bij de twee cursussen "Toegepaste Statistiek" 1) en "Elementaire Mathematische Statistiek" 2). Zij bestrijken slechts een zeer klein gedeelte van de gangbare statistische methoden, zoals uit de verderop gegeven formules en tabellen blijkt. De opgaven (of gedeelten daarvan), die aansluiten bij de eerstgenoemde cursus zijn van een sterretje voorzien.

Van alle oefeningen is een volledig uitgewerkte oplossing opgenomen, waarbij alleen van de in dit boekje opgenomen tabellen en formules gebruik gemaakt wordt. In sommige gevallen is naast een voor de hand liggende oplossing een tweede oplossingsmethode beschreven, die minder voor de hand ligt, maar sneller tot het doel leidt.

De volgorde der opgaven, hoewel niet aselect, is evenmin systematisch. Dit heeft ten doel de gewenste oplossingsmethode niet aan te geven door de plaats van de opgave.

NOTATIES

De in de statistiek gebruikelijke notaties en definities worden bekend ondersteld. Enkele daarvan vermelden wij hieronder:

wh = waarschijnlijkheid,
whn= waarschijnlijkheden,
whr= waarschijnlijkheidsrekening,
$\approx$ = bij benadering gelijk aan,
$\mathcal{E}$ = mathematische verwachting,
μ, σ^2 = verwachting, variantie,
k = tweezijdige overschrijdingskans,
k_r = rechtséénzijdige overschrijdingskans,
k_l = linkséénzijdige overschrijdingskans.

Stochastische grootheden worden door onderstreepte letters aangegeven. Dezelfde letters, niet onderstreept, stellen getallen of algebraïsche variabelen voor.

----

1) Rapport S 120 van het Mathematisch Centrum, Amsterdam, door Prof.Dr J. Hemelrijk en Ph. van Elteren.

2) Rapport S 95 van het Mathematisch Centrum, Amsterdam door Prof.Dr J. Hemelrijk.

$\underline{u}$ wordt steeds gebruikt voor een normaal verdeelde grootheid met gemiddelde 0 en spreiding 1 (een N(0,1) verdeelde grootheid).

$N(\mu,\sigma)$ geeft een normale verdeling met verwachting μ en spreiding σ aan.

FORMULES

1. De axioma's der whr, voor zover die bij deze opgaven expliciet te pas komen, luiden:

$$(1)\quad \begin{cases} 1)\ 0 \leqq P[A] \leqq 1 & \text{voor iedere gebeurtenis } A, \\ 2)\ P[A]=0 & \text{als } A \text{ onmogelijk is,} \\ 3)\ P[A]=1 & \text{als } A \text{ zeker is,} \\ 4)\ P[A \text{ of } B] = P[A]+P[B]-P[A \text{ en } B]. \end{cases}$$

2. Normale benadering. Bij een aantal opgaven moet gebruik gemaakt worden van de normale benadering voor de verdeling van een stochastische grootheid. Geven wij deze grootheid aan met $\underline{x}$ en is $\mathcal{E}\underline{x}=\mu$ en $\sigma^2\{\underline{x}\} = \sigma^2$, terwijl $\underline{u}$ een N(0,1) - verdeelde stochastische grootheid voorstelt, dan geldt:

$$(2)\qquad P[\underline{x} \geqq x] = P\left[\frac{\underline{x}-\mu}{\sigma} \geqq \frac{x-\mu}{\sigma}\right] \approx P\left[\underline{u} \geqq \frac{x-\mu}{\sigma}\right]$$

en deze laatste kan in tabel 1 van de N(0,1) - verdeling opgezocht worden.

Is $\underline{x}$ discreet verdeeld en zijn de afstanden tussen opeenvolgende mogelijke waarden van $\underline{x}$ gelijk aan 1, dan is het in de regel wenselijk een continuïteitscorrectie toe te passen, die tot de volgende formule leidt:

$$(2')\qquad P[\underline{x} \leqq x] \approx P\left[\underline{u} \leqq \frac{x-\mu+\frac{1}{2}}{\sigma}\right]$$

en

$$(2'')\qquad P[\underline{x} \geqq x] \approx P\left[\underline{u} \geqq \frac{x-\mu-\frac{1}{2}}{\sigma}\right].$$

3. Betrouwbaarheidsintervallen voor een onbekende fractie of wh

Zijn in een steekproef van n exemplaren uit een grote populatie x exemplaren gevonden met een zeker kenmerk A en stelt θ de (onbekende) fractie aan exemplaren met kenmerk A in deze populatie voor, dan geldt, behoudens een onbetrouwbaarheid α:

(3) $$\theta_* \leq \theta \leq \theta^*,$$

waarin

(4) $$\theta_* = \frac{x+\frac{1}{2}c - \sqrt{c\left(\frac{x(n-x)}{n} + \frac{1}{4}c\right)}}{n+c}$$

en

(5) $$\theta^* = \frac{x+\frac{1}{2}c + \sqrt{c\left(\frac{x(n-x)}{n} + \frac{1}{4}c\right)}}{n+c},$$

en waarin c , afhankelijk van α , in de onderstaande tabel opgezocht kan worden. (zie ook de opmerking op blz. 6).

Tabel A.

α (tweezijdig)	c
0,2	1,64
0,1	2,71
0,05	3,84
0,01	6,63

De éénzijdige uitspraken $\theta \leq \theta^*$ en $\theta_* \leq \theta$ gelden, ieder afzonderlijk, met onbetrouwbaarheidsdrempel $\frac{1}{2}\alpha$. De gelijkheidstekens kunnen naar willekeur weggelaten worden.

Deze formules berusten op de normale benadering van de binomiale verdeling en zijn dus niet exact. De benadering is goed, indien n niet te klein is en x niet te veel van $\frac{1}{2}n$ afwijkt. Is $x << n$ (of $n-x << n$), dan kan voor $x \leq 10$ (resp. $n-x \leq 10$) gebruik gemaakt worden van tabel 2, die op de Poissonbenadering van de binomiale verdeling berust.

Is $x >> \frac{1}{2}c$, $n >> c$ en $x(n-x) >> \frac{1}{4}nc$, dan kunnen de formules (4) en (5) vereenvoudigd worden tot :

(4') $$\theta_* = \frac{1}{n}\left(x - c'\sqrt{\frac{x(n-x)}{n}}\right)$$

en

(5') $$\theta^* = \frac{1}{n}\left(x + c'\sqrt{\frac{x(n-x)}{n}}\right),$$

waarin c' uit tabel B (afgeleid uit tabel 1) af te lezen is.

Tabel B.

α (tweezijdig)	c' ($=\sqrt{\bar{c}}$)
0,2	1,28
0,1	1,65
0,05	1,96
0,01	2,58

Deze methodes zijn ook toepasbaar als wij niet te maken hebben met een steekproef uit een eindige populatie, maar met n stochastisch onafhankelijke experimenten die ieder een kans θ op succes hebben; x stelt dan het aantal successen voor.

4. Toets voor de hypothese $p = p_0$

Gegeven zijn n stochastisch onafhankelijke experimenten met ieder een kans p op succes. Gevraagd wordt de hypothese $H_0: p=p_0$ te toetsen. Dit kan geschieden door een betrouwbaarheids-interval voor p te bepalen (zie 3) en H_0 te verwerpen als dit interval p_0 niet bevat. Is p niet te ver van $\frac{1}{2}$ verwijderd, dan kan men ook de overschrijdingskans, behorende bij het gevonden aantal (x) successen met behulp van de normale benadering bepalen. De linker- resp. rechteroverschrijdingskans vindt men dan door deze op te zoeken voor de waarde

(6) $$u = \frac{x - np + \frac{1}{2}}{\sqrt{npq}} \quad \text{resp.} \quad u = \frac{x - np - \frac{1}{2}}{\sqrt{npq}} .$$

De tweezijdige overschrijdingskans is gelijk aan het dubbele van de kleinste der beide éénzijdige overschrijdingskansen, of ook aan het dubbele van de rechteroverschrijdingskans, behorende bij

(7) $$u = \frac{|x - np| - \frac{1}{2}}{\sqrt{npq}} .$$

De kritieke waarden, voor gegeven tweezijdige onbetrouwbaarheidsdrempel α , zijn

(8) $$np \pm \left\{ c'\sqrt{npq} + \tfrac{1}{2} \right\},$$

waarin c' in tabel B opgezocht kan worden. Bij éénzijdige kritieke zones vindt men de grenzen eveneens op deze wijze, maar dan met een c' , die bij een α behoort, gelijk aan het dubbele van de éénzijdige α . De continuïteitscorrectie ($\pm\frac{1}{2}$ in (6), (7) en (8)) kan weggelaten worden, als n groot is.

De tekentoets is een speciaal geval van deze toets, dat optreedt, als $p_0 = \frac{1}{2}$ is.

5. Betrouwbaarheidsintervallen berustend op de verdeling van Student

Is $x_1, \ldots, x_n$ een steekproef uit een $N(\mu, \sigma)$-verdeling, dan geldt, behoudens een onbetrouwbaarheid α:

$$(9) \qquad \bar{x} - t_\alpha s/\sqrt{n} \leqq \mu \leqq \bar{x} + t_\alpha s/\sqrt{n},$$

waarin $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ en $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

en waarin t_α de waarde is, die in tabel 4 voor $\nu = n-1$ in de kolom met $k = \alpha$ gevonden wordt.

Zijn $x_1, \ldots, x_n$ en $y_1, \ldots, y_m$ twee stochastisch onafhankelijke steekproeven uit een $N(\mu_1, \sigma)$ en $N(\mu_2, \sigma)$- verdeling, dan geldt:

$$(10) \qquad \bar{x} - \bar{y} - t_\alpha \sqrt{S_1^2 + S_2^2}\sqrt{\frac{n+m}{nm(n+m-2)}} \leqq \mu_1 - \mu_2 \leqq \bar{x} - \bar{y} + t_\alpha \sqrt{S_1^2 + S_2^2}\sqrt{\frac{n+m}{nm(n+m-2)}},$$

waarin $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ en $S_1^2 = \overbrace{\Sigma(x_i - \bar{x})^2}^{= \Sigma x_i^2 - \frac{1}{n}(\Sigma x_i)^2}$ en analoog voor $\bar{y}$ en S_2. Hierin heeft t_α dezelfde betekenis als in (9), maar nu met $\nu = n+m-2$.

De gelijkheidstekens kunnen naar willekeur weggelaten worden. Voor éénzijdige betrouwbaarheidsgrenzen is α gelijk aan de helft van de in tabel 4 opgegeven waarden van k.

6. Toets van Wilcoxon met behulp van de normale benadering

Voor de toetsingsgrootheid $\underline{u}$ van Wilcoxon geldt:

$$(11) \qquad \mu = \tfrac{1}{2} nm$$

en

$$(12) \qquad \sigma^2 = \frac{1}{12} mn(m+n+1) - \frac{1}{12}\frac{mn}{(m+n)(m+n-1)}\Sigma(t_i^3 - t_i),$$

waarin m en n de uitgebreidheden der twee steekproeven voorstellen en t_i de uitgebreidheden der groepen gelijke waarnemingen in de beide steekproeven tezamen genomen.

Een andere vorm voor de variantie van $\underline{U}$ is:

$$(12') \qquad \sigma^2 = \frac{1}{12}\,\frac{mn\{(m+n)^3 - (1^3 g_1 + 2^3 g_2 + 3^3 g_3 + \dots)\}}{(m+n)(m+n-1)}$$

waarin $g_j\,(j=1,2,\dots)$ het aantal groepen van j gelijke waarnemingen voorstelt (m.a.w. g_j is het aantal der t_i die gelijk aan j zijn).

De kritieke waarden zijn nu

$$(13) \qquad U_\ell = \mu - c'\sigma - \tfrac{1}{2} \qquad \text{en} \qquad U_r = \mu + c'\sigma + \tfrac{1}{2}$$

waarin μ en σ de door (11) en (12) aangegeven waarden bezitten en c' in tabel B opgezocht kan worden.

De overschrijdingskansen worden berekend (zie 2) met behulp van

$$(14) \qquad u = \frac{U - \mu + \frac{1}{2}}{\sigma}\,, \qquad u = \frac{U - \mu - \frac{1}{2}}{\sigma} \qquad \text{resp.} \qquad u = \frac{|U - \mu| - \frac{1}{2}}{\sigma}\,.$$

7. Binomiale verdeling.

De exacte formule der binomiale verdeling luidt:

$$(15) \qquad P[\underline{x} = x] = \binom{n}{x} p^x q^{n-x}\,,$$

met

$$(16) \qquad \mathcal{E}\underline{x} = \mu = np\,; \quad \sigma^2\{\underline{x}\} = npq\,.$$

Hierin stelt n het aantal onafhankelijke experimenten voor, p de kans op succes bij ieder hiervan en x het aantal successen. Met behulp van (16) kan men, als n niet te klein is en p niet te ver van $\frac{1}{2}$ afwijkt, de normale benadering toepassen volgens de onder 2 beschreven methode.

Opmerkingen bij formules (4) en (5) op blz. 3. naar voren en bij de oplossingen toepassen

Bij de formules (4) en (5) is geen continuïteitscorrectie in rekening gebracht. Dit kan geschieden door in (4) x te vervangen door $x - \frac{1}{2}$ en in (5) door $x + \frac{1}{2}$. Bij de oplossingen is deze verfijning niet toegepast.

OPGAVEN

1.* Een consument koopt een partij goederen, die volgens de verzekering van de fabrikant niet meer dan 2% aan ondeugdelijke exemplaren bevat.

Om dit te controleren neemt de koper een steekproef van 90 exemplaren waaronder hij 4 defecten vindt. Hij wenst slechts een kans van hoogstens 0,05 te lopen om ten onrechte te reclameren. Moet hij dan reclameren op grond van het gevonden aantal van 4 defecte exemplaren?

Hoeveel defecte exemplaren moet hij in de steekproef vinden om tot reclameren over te gaan?

2.* Een chemisch analyste moet een aantal titraties in duplo verrichten. De te titreren oplossing wordt daartoe, direct na de bereiding in twee kolfjes gedaan en terwijl de eerste titratie plaats vindt staat het tweede kolfje te wachten.

Zij verkrijgt de volgende uitkomsten.

Oplossing nr	1e titratie	2e titratie
1	21,24	25,83
2	16,84	17,35
3	15,52	16 12
4	25,68	28,54
5	24,04	24,58
6	19,77	27,42
7	11,92	14,73
8	28,83	27,52
9	17,38	14,91
10	11,01	19,87
11	23,43	24,38
12	17,16	20,55

Ga na (met onbetrouwbaarheidsdrempel 0,05) of het wachten invloed heeft uitgeoefend op de uitkomst van de titraties.

3.* Een onderzoeker wil uit een school met 6 klassen een aselecte steekproef van 10 kinderen nemen. De klassen bevatten verschillende aantallen kinderen:

klasse	1	2	3	4	5	6
aantal kinderen	20	25	15	10	10	15

Als hulpmiddel voor het nemen van deze steekproef is een tabel van aselecte getallen beschikbaar (tabel 6). Neem met behulp daarvan een steekproef van 10 kinderen; beschrijf hoe U dit doet, wat de gebruikte aselecte getallen zijn en hoeveel kinderen uit ieder der klassen in de steekproef voorkomen.

4*. Bij 225 onafhankelijke experimenten, die alle dezelfde kans p op succes hebben, worden 85 successen gevonden. Toets de hypothese, dat $p = \frac{1}{2}$ is met $\alpha = 0{,}05$. Toets eveneens met $\alpha = 0{,}05$ de hypothese dat $p = 1/3$. Voer beide toetsen tweezijdig en linkséénzijdig uit.

5*. Over het aantal verkeersongevallen met dodelijke afloop in twee steden A en B zijn de volgende gegevens beschikbaar:

Jaar	in A	in B
1920	125	103
1921	110	97
1922	89	93
1923	130	80
1924	145	110
1925	137	115
1926	105	93
1927	131	87
1928	126	111
1929	117	102
1930	142	81
1931	97	101
1932	99	105
1933	109	75
1934	116	83
1935	105	92
1936	121	107

Toets, of er in één van beide steden systematisch meer ongevallen met dodelijke afloop optreden dan in de andere (onbetrouwbaarheidsdrempel 0,05). Wat is Uw conclusie? Welke verklaringen van het door U geconstateerde verschijnsel acht U aannemelijk?

6*. Een fabrikant van "grijpautomaten" (miniatuurhijskranen, die men soms op de kermis ziet en waarmee men kan trachten een prijs te winnen) beweert, dat het spelen met deze apparaten op behendigheid berust en niet een zuiver kansspel is. Om dit te bewijzen laat hij eerst een groot aantal ongeroutineerde spelers de machine bedienen. Daarbij blijkt dat de kans - voor een ongeroutineerde speler - om een prijs te krijgen ongeveer gelijk is aan 1/50. Vervolgens speelt hij zelf 75 maal en haalt daarbij 8 prijzen.

a) Ga na of zijn bewering juist is (onbetrouwbaarheidsdrempel 0,05).

b) Ga ook na hoeveel prijzen hij bij 75 maal proberen minstens moet halen om zijn bewering te staven (met dezelfde onbetrouwbaarheidsdrempel).

7 $\underline{x}$ is binomiaal verdeeld met $n = 6$ en $p = \frac{1}{3}$.
Bereken de voorwaardelijke verdeling van $\underline{x}$ onder de voorwaarde $\underline{x} \leq 3$ en gemiddelde en spreiding van deze verdeling.

8. a) $\underline{u}$ is een $N(0,1)$ verdeelde stochastische grootheid. Bepaal:

$$P[\underline{u} \leq 1{,}26]$$
$$P[0{,}13 < \underline{u} \leq 1{,}26]$$
en $$P[\underline{u} \geq 0{,}9].$$

b) $\underline{x}$ is normaal verdeeld met gemiddelde 2,1 en spreiding 0,7.
Bepaal:

$$P[|\underline{x}| \geq 0{,}9].$$

c) $\underline{y}$ is binomiaal verdeeld met $n = 25$ en $p = 0{,}25$. Bepaal met behulp van de normale benadering met continuïteitscorrectie, de kansen

$$P[5 < \underline{y} \leq 12]$$
$$P[5 \leq \underline{y} < 12].$$

9. $\underline{x}$ is een stochastische grootheid, die de waarden $a, 2a, 3a, \ldots, ka$ aanneemt met gelijke waarschijnlijkheden.

a) Bereken gemiddelde en variantie van $\underline{x}$.

b) Bereken de wh-verdeling van $1/\underline{x}$ en gemiddelde en variantie daarvan voor $k=5$.

10*. Twee verschillende katoenen weefsels werden, onder overigens gelijke omstandigheden, op hun slijtbestendigheid beproefd. Van beide weefsels verrichtte men 10 slijtageproeven, waarbij de volgende uitkomsten verkregen werden (beide uitgedrukt in dezelfde eenheid).

Aantal toeren voor breuk

Weefsel A	Weefsel B
629	652
713	654
681	731
736	728
764	712
646	783
654	663
657	688
697	717
654	676

Onderzoek ~~Toets~~ of één van beide weefsels slijtbestendiger is dan het andere ($\alpha = 0{,}05$). (Toets de nulhypothese: beide weefsels even sterk).

11.* a) Hoe luidt de definitie van stochastische onafhankelijkheid van twee gebeurtenissen A en B?

b) Beschouw het volgende geval: uit een pak van 100 kaarten, die de nummers 1,.....,100 dragen, wordt aselect één kaart getrokken. Laat gebeurtenis A zijn: "het getrokken nummer is ≤ 40" en gebeurtenis B: "het getrokken nummer is ≥ 20". Ga na, of A en B stochastisch onafhankelijk zijn of niet.

c) Als wij twee gebeurtenissen "positief afhankelijk" noemen als zij vaker tezamen optreden dan bij onafhankelijkheid het geval zou zijn en "negatief afhankelijk" als dit minder vaak het geval is, ga dan na voor welke waarden van n de gebeurtenis A en de gebeurtenis C: "het getrokken nummer is $\geq n$", positief resp. negatief afhankelijk zijn.

12.* In een vijver wil men de visbevolking schatten. Men vangt daartoe 50 vissen welke van een merkteken worden voorzien.

Vervolgens laat men deze vissen weer zwemmen en na verloop van enige tijd vangt men 100 vissen. Hierbij blijken 9 gemerkte vissen te zijn.

Geef grenzen aan, waartussen de visbevolking, behoudens een onbetrouwbaarheid 0,05 zal liggen, in de onderstelling, dat de vispopulatie niet veranderd is en dat de vangsten aselect waren.

13. a) Als $\underline{x}$ een stochastische grootheid is, die de waarden 1,2,....,100 met gelijke waarschijnlijkheid aanneemt, bereken dan de spreiding van $\underline{x}$.

b) Laat $\underline{y}$ het aantal onafhankelijke waarnemingen van $\underline{x}$ voorstellen tot voor het eerst de waarde $\underline{x} \geqq 80$ gevonden wordt. Bereken $P[\underline{y} \leqq 3]$.

14.* Een accountant heeft gegevens verzameld over de verkoop van een bepaald artikel (b.v. gasfornuizen) door 2 winkels met ogenschijnlijk ongeveer even grote omzet.

De aantallen verkochte gasfornuizen zijn:

Jaar	winkel A	winkel B
1937	209	163
1938	218	204
1939	182	183
1940	200	238
1941	197	203
1942	232	273
1946	232	274
1947	253	264
1948	277	227
1949	227	285
1950	284	244
1951	222	285
1952	298	290

Ga na, of één der beide winkels een grotere verkoop van gasfornuizen heeft dan de andere. (met onbetrouwbaarheidsdrempel $\alpha = 0,05$).

15.* Uit een pak van N kaarten, die van 1 tot N genummerd zijn, wordt aselect één kaart getrokken. A stelt de gebeurtenis "het getrokken nummer is $\leqq a$" voor en B de gebeurtenis "het getrokken nummer is $\geqq b$" (a en b beide $\geqq 0$ en $\leqq N$).

Ga na, of er getallen voor N,a en b te vinden zijn, waarvoor de gebeurtenissen A en B stochastisch onafhankelijk zijn.

16. Twee personen kiezen ieder aselect 10 verschillende van de getallen van 1 tot 100. Wat is de kans op minstens één paar gelijken?

17.* a) Twee soorten gloeilampen moeten met elkaar vergeleken worden wat hun levensduur betreft. Daartoe worden van de eerste soort 10 exemplaren en van de tweede 12 exemplaren genomen en deze worden gebrand tot zij doorbranden. De uitkomsten zijn (in uren):

A: 625, 637, 710, 770, 820, 843, 856, 920, 1070, 1225.

B: 630, 683, 780, 830, 889, 970, 1028, 1150, 1210, 1470, 1520, 2090.

Onderzoek ~~Toets~~, of één van beide soorten een systematisch langere levensduur heeft dan de andere soort (onbetrouwbaarheidsdrempel 0,05).

b) Indien men de boven beschreven proef wil bekorten tot maximaal 1000 uur nemen de gegevens de volgende vorm aan:

A: 625, 637, 710, 770, 820, 843, 856, 920, 2 waarnemingen > 1000.

B: 630, 683, 780, 830, 889, 970, 6 waarnemingen > 1000.

Toets ook voor deze gegevens de hypothese, die U in deel a) van deze opgave getoetst hebt. (Aanwijzing: beschouw de waarnemingen, die > 1000 zijn, alle als gelijk uitgevallen waarnemingen).

18. Bij een rij onafhankelijke experimenten heeft ieder experiment een kans p op succes en q op mislukking. Als $\underline{n}$ het aantal experimenten is tot en met het eerste succes, bereken dan

$$P[\underline{n}=n] \quad \text{en} \quad \mathcal{E}\underline{n} \quad \text{en} \quad \sigma^2\{\underline{n}\}.$$

19.* Men neemt van een partij goederen, die volgens de bewering van de fabrikant niet meer dan 1% ondeugdelijke exemplaren bevat, een steekproef van 150 exemplaren. Hierin vindt men 5 ondeugdelijke exemplaren.

Indien men nu alleen wil reclameren, indien uit de steekproef blijkt, dat de partij goederen meer dan 1% ondeugdelijke exemplaren bevat en men bij een partij die goed is, hoogstens een kans 0,05 wil hebben om toch te reclameren, zal men dan op grond van dit steekproefresultaat reclameren of niet?

Hoe groot mag het aantal ondeugdelijke exemplaren in de steekproef van 150 exemplaren bij gebruik van dezelfde onbetrouwbaarheidsdrempel maximaal zijn, opdat men niet zal reclameren?

20. a) Wat is de definitie van het onderscheidingsvermogen van een toets?

b) Bereken voor een aantal waarden van p het onderscheidingsvermogen voor de rechter-éénzijdige tekentoets met $n=5$ en met $x=5$ als kritieke zone en maak een schets van dit onderscheidingsvermogen.

21*. Vier voorwerpen $v_1, \ldots, v_4$ worden op de volgende wijze in 2 groepen A en B van ieder 2 voorwerpen verdeeld: in de volgorde $v_1, \ldots, v_4$ wordt voor ieder voorwerp geloot met een zuivere munt, of het in A of in B komt. Dit wordt voortgezet tot ofwel A ofwel B twee voorwerpen bevat, waarna het restant in de andere groep wordt ingedeeld.

Ga na: a) of ieder voorwerp gelijke kans heeft om in A of B te komen?

b) of de verdeling aselect is in die zin, dat alle tweetallen (v_i, v_j) gelijke kans hebben on tezamen groep A te vormen?

22. Gegeven zijn vier voorwerpen $v_1, \ldots, v_4$ en een zuivere munt.

Gevraagd: a*) Deze vier voorwerpen aselect in 2 groepen van 2 voorwerpen te verdelen, dus in die zin, dat alle tweetallen $v_i v_j$ gelijke kans hebben om tezamen één groep te vormen.

b) Bereken de verwachting van het aantal worpen met de munt, dat uitgevoerd moet worden, om deze verdeling tot stand te brengen.

23. a) Een speler A verricht een reeks onafhankelijke worpen met een zuivere munt. Zij $\underline{n}$ het aantal worpen, dat hij verrichten moet om voor het eerst kruis te verkrijgen. Bereken

$$P[\underline{n}=1],\ P[\underline{n}=2], \ldots,\ P[\underline{n}=n]$$

en

$$\mathcal{E}\underline{n} \quad \text{en} \quad \sigma^2\{\underline{n}\}.$$

b) Een tweede speler, B, doet hetzelfde als A. Noem voor deze speler het aantal worpen tot en met de eerste maal kruis $\underline{m}$. Bereken de kans

$$P[\underline{n}=\underline{m}].$$

c) Als A van B wint, indien hij het eerst kruis krijgt, en B van A, indien dit hem gelukt, terwijl zij remise spelen als zij tegelijkertijd kruis verkrijgen (dus als $\underline{m}=\underline{n}$ is), hoe groot is dan de winstkans van A en hoe groot is die van B?

24. $\underline{x}$ bezit een discrete verdeling. Bewijs dat de som van de linker- en rechteroverschrijdingskans, behorende bij een door $\underline{x}$ aangenomen waarde groter dan 1 is.

Geldt dit ook voor continu verdeelde grootheden?

25. $\underline{x}$ en $\underline{y}$ zijn onderling onafhankelijk discreet verdeelde stochastische grootheden. Bewijs dat ook de grootheden $\underline{x}^2$ en $\underline{y}$ onderling onafhankelijk verdeeld zijn.

26*. a) Met een onzuivere munt worden tweetallen onafhankelijke worpen uitgevoerd tot er een tweetal optreedt, waarbij éénmaal kruis en éénmaal munt optreedt. Er zijn dan twee mogelijke volgorden KM en MK. Bewijs dat deze beide een wh $\frac{1}{2}$ hebben.

b) Hoe kan men met behulp van een onzuivere munt een groep van 6 voorwerpen zodanig in twee groepen A en B van ieder 3 voorwerpen verdelen dat ieder der 6 voorwerpen een kans $\frac{1}{2}$ bezit om in groep A terecht te komen.

27. Bij een statistisch onderzoek worden op vijf groepen waarnemingen, die onderling onafhankelijk verkregen zijn, statistische analyses toegepast, die tot de conclusie A,B,C,D en E leiden. Voor ieder der analyses geldt, dat de bijbehorende conclusie behoudens een onbetrouwbaarheid α (die voor alle vijf dezelfde is) geldt.

Beantwoordt de volgende vragen:

a) Hoe groot is bij een dergelijke werkwijze de kans, dat er onder de conclusies één of meer foute voorkomen.
b) Hoe groot is de kans dat ze allemaal fout zijn?
c) Welke numerieke waarden nemen deze kansen aan voor $\alpha = 0{,}02$
d) Hoe groot moet men α nemen om er, behoudens een onbetrouwbaarheid 0,05, op te kunnen rekenen, dat alle vijf conclusies juist zijn.

28*. Een fabrikant, die een massaproduct in grote partijen verkoopt garandeert aan zijn afnemers, dat deze partijen niet meer dan 3% aan defecte exemplaren bevatten.
Om aan deze garantie redelijk goed te blijven voldoen, keurt hij uit iedere partij een steekproef van exemplaren.
Op grond van de keuring van deze steekproef keurt hij de gehele partij, waaruit de steekproef afkomstig is goed of af.
Van die partijen die 3% of meer aan defecte exemplaren bevatten wil hij er hoogstens één op de 10 doorlaten(d.w.z per ongeluk toch goedkeuren).
Beantwoord de volgende vragen:

a) Als hij steekproeven van 100 exemplaren neemt, bij welk aantal defecte exemplaren in de steekproef moet hij de partij dan afkeuren?
b) Als hij partijen wil goedkeuren, wanneer in de steek-

proef 0 of 1 defect exemplaar voorkomt en afkeuren als het er twee of meer zijn, hoe groot moet hij dan n minstens nemen?

29*. Uit een pot met erwten, waarvan een onbekende fractie θ rood is, wordt een steekproef van n erwten getrokken, waaronder x rode voorkomen.

Wat is de lengte van het bijbehorende tweezijdige betrouwbaarheidsinterval voor θ, indien men $\alpha = 0{,}01$ neemt (gebruik de eenvoudigste benaderingsformule). Wat is de invloed van de vergroting van n, bij gelijkblijvende α en θ? Als $\theta = 0{,}4$ en men wenst een betrouwbaarheidsinterval (met $\alpha = 0{,}01$) van ongeveer de lengte $0{,}1$ te vinden, hoe groot zal men dan n ongeveer moeten nemen?

30*. a) Uit een partij goederen van 10.000 exemplaren wordt een steekproef van 40 exemplaren genomen. Ieder exemplaar wordt goed- of afgekeurd; het aantal afgekeurde exemplaren wordt geteld en op grond van dit aantal wordt een betrouwbaarheidsinterval bepaald voor het aantal ondeugdelijke exemplaren in de partij. Het werkelijke aantal ondeugdelijke exemplaren in de partij is 500 Hoe groot is de kans, dat een willekeurig getrokken exemplaar ondeugdelijk is?

b) Boots, op grond van dit gegeven, het nemen van een steekproef van 40 exemplaren na met behulp van een tabel van aselecte getallen. Beschrijf, hoe U dit doet en tel in de door U nagebootste steekproef het aantal "defecte" exemplaren. Bepaal op grond van het door U gevonden aantal een tweezijdig betrouwbaarheidsinterval (met onbetrouwbaarheidsdrempel 0,05) voor het aantal ondeugdelijke exemplaren in de partij en ga na, of dit interval het werkelijke aantal daarvan bevat of niet.

31. Een organisator van een vuurwerk weet uit ervaring, dat van de vuurpijlen, die hij gewoonlijk gebruikt, ongeveer één op de 5 weigert. In verband daarmee maakt hij een aantal reserve pijlen gereed. Hij wenst het publiek minstens 20 goede vuurpijlen te laten zien en hij aanvaardt hoogstens een kans van 1/100, dat hem dit niet zal lukken.

Hoeveel reserve pijlen moet hij opstellen?

32. Een fabrikant wil een zeer grote partij goederen keuren op het percentage defecte exemplaren. Hij wil deze partij liever

niet afkeuren indien er niet meer dan 2% aan defecte exemplaren in voorkomen.
Ter keuring neemt hij een steekproef van 150 exemplaren en daaronder vindt hij 5 defecten. Hij vermoedt dat dit resultaat hem niet tot afkeuren zal leiden, indien hij hoogstens een kans $\alpha = 0{,}05$ op ten onrechte afkeuring toelaat.

a)* Is zijn vermoeden juist?

b)* Bij welk aantal defecten onder 150 exemplaren zal hij tot afkeuring overgaan?

c) Anderzijds is de fabrikant er niet zeker van, dat zijn steekproef voldoende groot is geweest. Hij zou de partij graag afkeuren indien deze 10% of meer aan defecten bevat en hij vraagt zich af, hoe groot de kans op afkeuring is indien het percentage defecten inderdaad gelijk is aan 10. Bereken een benadering van deze kans met behulp van de normale verdeling.

33. a) A en B spelen een spel, waarbij, naar men zegt, de winstkansen voor beide gelijk zijn. Om na te gaan of dit zo is, spreken zij af, dat zij 100 maal zullen spelen. Indien A daarbij 65 maal wint en B 35 maal, is er dan reden om aan de gelijkheid der winstkansen te twijfelen? Toetst U in dit geval één- of tweezijdig?

b) Indien men bij deze toets een onbetrouwbaarheidsdrempel 0,05 gebruikt, hoe vaak mag A dan hoogstens winnen zonder dat de hypothese van de gelijkheid der winstkansen verworpen moet worden?

c) Hoe groot moet p (de kans op winst voor A) ongeveer zijn om de toets met onbetrouwbaarheidsdrempel 0,05 een onderscheidingsvermogen 0,8 te geven? De continuïteitscorrectie mag bij deze berekening weggelaten worden.

34. Twee gebeurtenissen A en B treden steeds tegelijk op. Bewijs met behulp van de axiomas van de waarschijnlijkheidsrekening dat:

$$P[A] = P[B].$$

35. Van twee gebeurtenissen A en B treedt B altijd op als A optreedt, maar niet noodzakelijk andersom. Bewijs met behulp van de axiomas van de waarschijnlijkheidsrekening:

a) $P[A \text{ en } B] = P[A]$
b) $P[A] \leq P[B]$
c) $P[A \text{ of } B] = P[B]$
d) $P[B|A] = 1$.

36.* a) Twee gebeurtenissen A en B sluiten elkaar uit. Zijn zij stochastisch onafhankelijk of niet?
b) Van twee gebeurtenissen A en B treedt B altijd op als A optreedt. Zijn zij stochastisch onafhankelijk of niet?
c) Van twee gebeurtenissen A en B treedt B zeker op. Zijn zij stochastisch onafhankelijk of niet?

37.* n voorwerpen moeten aselect verdeeld worden in twee groepen A en B van k resp. n-k voorwerpen. Alle k-tallen moeten dus gelijke wh bezitten om groep A te vormen.

Hoe kan dit uitgevoerd worden:
a) met behulp van een zuivere munt?
b) met behulp van de tabel van de aselecte getallen?

38. Men verricht n onafhankelijke experimenten, waarbij de kans op succes voor ieder experiment $p=0{,}2$ is.

Hoe groot moet men n minstens nemen om, met de tweezijdige tekentoets met onbetrouwbaarheid $\alpha = 0{,}05$, behoudens een waarschijnlijkheid 0,1 de afwijking van $p=\frac{1}{2}$ te constateren?
(De continuïteitscorrectie behoeft niet te worden toegepast).

39. Een snoepgoed-winkelier ontvangt van zijn leverancier een kist met "surprise-caramels". Volgens opgave van de leverancier is bij 20% van de caramels een bonnetje in de wikkel verpakt, dat recht geeft op een gratis "surprise-caramel".

Aannemende dat dit percentage juist is en dat de bonnetjes aselect over de caramels verdeeld zijn, wordt gevraagd:
a) hoe groot is de kans, dat een jongetje, dat 5 caramels koopt, er meer dan 5 krijgt?
b) hoe groot is de kans dat hij er precies 6 krijgt?
c) hoe groot is de kans dat hij er precies 7 krijgt?
d) hoe groot is de kans dat hij er meer dan 7 krijgt?

40. De winkelier uit het vorige vraagstuk wil er, in verband met de aan deze surprise-actie verbonden kosten redelijk zeker van zijn dat het aantal caramels met een "prijs" niet hoger is

dan 20%. Hij kan echter niet alle caramels open maken, daar hij de wikkels dan niet netjes meer dicht kan vouwen, hetgeen wantrouwen opwekt bij zijn jeugdige klanten. Hij besluit daarom tot het nemen van een steekproef van 100 exemplaren.
Vragen:

a) Indien het opgegeven percentage aan prijzen juist is, wat zijn dan gemiddelde en spreiding van het aantal prijzen, dat hij in zijn steekproef vindt.

b) Indien de winkelier een onbetrouwbaarheidsdrempel $\alpha = 0{,}05$ toelaat, wanneer zal hij dan besluiten bij zijn leverancier te protesteren, dat de opgave van het percentage prijzen niet juist is geweest en dat hem daardoor schade is berokkend. (Gebruik hierbij de normale benadering).

41 Een der jeugdige klanten van deze winkelier klaagt bij zijn vader, die statisticus is, dat hij nu al drie weken lang al zijn zakgeld in caramels belegd heeft, maar nog steeds geen enkele "prijs" gewonnen heeft. Hij vermoedt, dat het in die winkel niet eerlijk toegaat. Zijn vader rekent uit, dat de kans op dit ongunstige resultaat van zijn zoontje, als het in die winkel wel eerlijk toegaat, ongeveer 0,07 is (om precies te zijn 0,0687) en hij deelt zijn zoontje mee dat hij wel een pechvogel is geweest, maar dat het nog te vroeg is om ruzie met de winkelier te gaan maken. Om het kind te troosten belooft hij hem dat zij samen naar de winkel zullen gaan en zoveel caramels zullen kopen dat er een kans van minstens 0,99 bestaat dat er minstens één prijs bij is.

Vragen:

a) Hoeveel caramels kan het jongetje van één week zakgeld kopen?

b) Hoeveel caramels moet de vader kopen om aan zijn belofte te voldoen, als het percentage inderdaad 20% is?

42. Gegeven, dat de gebeurtenissen A en B stochastisch onafhankelijk zijn.

Bewijs dat ook A en $\bar{B}$ waarin $\bar{B}$ de gebeurtenis " B treedt niet op" voorstelt, stochastisch onafhankelijk zijn.

43. Twee zuivere dobbelstenen worden, onafhankelijk van elkaar geworpen. Het aantal ogen dat de éne geeft, noemen wij $\underline{x}$, het

aantal ogen van de andere $\underline{y}$. Verder is $\underline{z}=\underline{x}+\underline{y}$ het aantal ogen, dat beide tezamen geven.

a) Zijn $\underline{x}$ en $\underline{z}$ stochastisch onafhankelijk of niet? Geef een bewijs voor de juistheid van Uw antwoord

Wij beschouwen nu twee mogelijke gebeurtenissen (bij één worp van beiden stenen), nl.:

A: $\underline{x}$ neemt een even waarde aan,

B: $\underline{z}$ neemt een oneven waarde aan.

b) Zijn A en B stochastisch onafhankelijke gebeurtenissen of niet? Bewijs Uw antwoord.

c) Bereken de voorwaardelijke waarschijnlijkheidsverdeling van $\underline{z}$ onder de voorwaarde dat $\underline{x}$ oneven is. Bereken van deze verdeling het gemiddelde en de spreiding. Bereken deze parameters ook voor de onvoorwaardelijke verdeling van $\underline{z}$ (waarbij dus aan $\underline{x}$ geen voorwaarde opgelegd wordt) en vergelijk de corresponderende parameters met elkaar. Wat merkt U hierbij op?

44. Onderstaande twee steekproeven zijn afkomstig uit twee normale verdelingen met gelijke spreiding

0,87	0,52
0,75	1,94
0,03	0,53
0,17	1,02
0,67	0,94
1,92	0,06
0,23	0,25
2,14	2,05
2,69	1,93
1,31	1,17
	1,38
	1,43
	1,74

a) Toets, met behulp van de toets van Student, de hypothese dat de gemiddelden van de beide verdelingen gelijk zijn

b) Bepaal met behulp van de verdeling van Student een tweezijdig betrouwbaarheidsinterval (met onbetrouwbaarheidsdrempel 0,05) voor het verschil van deze twee gemiddelden.

c) Pas op de twee steekproeven de toets van Wilcoxon toe

45. Uit een vaas met n witte en m zwarte ballen worden zonder teruglegging r ballen aselect getrokken.

a) Noem het aantal witte ballen onder de r getrokken ballen $\underline{a}$ en bewijs dat de wh-verdeling van $\underline{a}$ gegeven wordt door:

$$P[\underline{a}=a] = \frac{\binom{n}{a}\binom{m}{b}}{\binom{N}{z}},$$

waarin $a+b=z$ en $n+m=N$.

b) Bereken $\mathcal{E}\underline{a}$, $\mathcal{E}\underline{a}^2$ en $\sigma^2\{\underline{a}\}$.

Opmerking: Bij afspraak is $\binom{a}{b}=0$ indien b negatief of $>a$ is.

46. Uit een vaas met n rode en m witte ballen worden zonder teruglegging k ballen getrokken en opzijgelegd, zonder dat naar de kleur wordt gekeken. Hoe groot is de kans dat de $(k+1)$ste bal wit is?

47. Twee stochastische grootheden $\underline{x}$ en $\underline{y}$ zijn beide binomiaal verdeeld; $\underline{x}$ met parameters $n=3$ en $p=\frac{1}{3}$, $\underline{y}$ met parameters $n=4$ en $p=\frac{1}{4}$ Verder is gegeven dat $\underline{x}$ en $\underline{y}$ onafhankelijk verdeeld zijn.

Bepaal de waarschijnlijkheidsverdeling van $\underline{z}=\underline{x}+\underline{y}$ en bereken gemiddelde en spreiding van deze verdeling.

48 Door een bepaalde wijze van prepareren (methode A) hoopt een textielfabrikant zijn garens sterker te maken dan zij zonder toepassing daarvan zijn. Om dit te onderzoeken, vervaardigt hij van twee soorten garens, die hij gewoonlijk in grote hoeveelheden produceert een partij met en zonder toepassing van methode A. Vervolgens past hij trekproeven op de zo verkregen garens toe en hij vindt de volgende numerieke resultaten (uitgedrukt in de een of andere sterkte eenheid).

Garensoort I	Garensoort II	
13,37	11,46	zonder A
12,66	10,69	
12,10	10,19	
13,20	10,34	
11,40	11,79	
12,97	11,52	met A
1[illegible],62	12,63	
12,54	10,59	
13,41	11,64	
13,57	11,70	

Ga, met behulp van een toets, met onbetrouwbaarheid 0,05 na:

a) Of de garensoorten I en II verschillend in sterkte zijn.

b) Of de fabrikant er wijs aan doet methode A in te voeren of niet.

49. De stochastische grootheid $\underline{x}$ bezit een binomiale verdeling met parameters n en p.

Bereken $\mathcal{E}\,\frac{1}{k}\sum_{i=1}^{k}(\underline{x}_i-\overline{\underline{x}})^2$, als $\underline{x}_1,\ldots,\underline{x}_k$ onderling onafhankelijk verdeeld zijn volgens dezelfde binomiale verdeling als $\underline{x}$, terwijl $\overline{\underline{x}}=\frac{1}{k}\sum_{i=1}^{k}\underline{x}_i$ is.

50. Aan een tentamen werd deelgenomen door studenten van twee richtingen A en B.
Van de 28 A-studenten zakten er 11 en slaagden er 17.
Van de 68 B-studenten zakten er 16 en slaagden er 52.

Indien wij aan de gezakten het kenmerk 0 en aan de geslaagden het kenmerk 1 toevoegen, kunnen wij beide groepen van 28 resp. 68 waarnemingen als een steekproef beschouwen. Passen wij op deze twee steekproeven de toets van Wilcoxon toe, dan kunnen wij toetsen of de kans op slagen voor studenten van beide richtingen gelijk was. ~~of niet~~.

Voer dit uit en formuleer Uw conclusie.

OPLOSSINGEN

1. Indien de koper op grond van het steekproefresultaat zou beslissen dat, behoudens een onbetrouwbaarheid 0,05, de partij goederen meer dan 2% defecte exemplaren bevat, dan zal hij reclameren. Dit is het geval als de ondergrens van het betrouwbaarheidsinterval voor het percentage defecte exemplaren hoger is dan 2%. In tabel 2 voor betrouwbaarheidsintervallen voor θ vinden wij voor $x = 4$ en $\alpha = 0{,}05$ dat $n\theta_* = 1{,}4$.
De ondergrens is dus, daar $n = 90$ is, $\frac{1{,}4}{90} \cdot 100\% = 1{,}56\%$ en daar deze lager is dan 2% zal de koper niet reclameren.

Bij welk aantal defecte exemplaren in een steekproef van 90 exemplaren zal hij wel reclameren?

De ondergrens van het betrouwbaarheidsinterval moet groter zijn dan 2%. Dus

$$\theta_* > 0{,}02$$
$$90\,\theta_* > 1{,}8.$$

Bij opzoeken in tabel 2 zien wij dat voor $x = 5$ de waarde 2,0 voor $n\theta_*$ gevonden wordt bij $\alpha = 0{,}05$, terwijl voor $x = 4$ (zie boven) de waarde 1,8 nog niet overschreden wordt.

Het kleinste aantal defecte exemplaren in een steekproef van 90 exemplaren, waarbij de koper zal reclameren bedraagt dus 5.

2. Indien het wachten van de oplossing geen invloed heeft op de uitkomst van de titratie, is (na weglating van eventuele gelijken) de kans, dat de 2° titratie hoger dan de 1° uitvalt, gelijk aan $\frac{1}{2}$ en omgekeerd.

De hypothese H_0: het wachten heeft geen invloed op de uitkomst van de titratie kunnen wij dus toetsen door de tweezijdige tekentoets toe te passen op de verschillen van de uitkomsten van de 2° en de 1° titratie. Wij toetsen tweezijdig omdat zowel de 2° titratie lager, als hoger gevonden kan worden.

Het resultaat is:

nummer oplossing	2° titr. - 1° titr.	nummer oplossing	2° titr. - 1° titr.
1	+	7	+
2	+	8	-
3	+	9	-
4	+	10	+
5	+	11	+
6	+	12	+

Wij nemen als toetsingsgrootheid $\underline{x}$ het aantal mintekens. Verder is $n=12$ en $H_0: p=\frac{1}{2}$,
dus

$$u=\frac{|x-np|-\frac{1}{2}}{\sqrt{npq}}=\frac{|2-6|-\frac{1}{2}}{1{,}732}=2{,}02.$$

De tweezijdige overschrijdingskans bedraagt 0,04. (tabel 1)

Conclusie:

Behoudens een onbetrouwbaarheid 0,05 verwerpen wij de hypothese, dat het wachten geen invloed heeft op de uitkomst van de titratie, ten gunste van de hypothese dat het wachten een zodanige invloed heeft dat voor de tweede titratie systematisch meer titratie vloeistof verbruikt wordt dan voor de eerste. Hetzelfde resultaat wordt direct verkregen met behulp van tabel 3.

3. De kinderen worden doorlopend genummerd van 1 tot 95.

klas	1	2	3	4	5	6
aantal kinderen	20	25	15	10	10	15
nummers	1t/m20	21t/m45	46t/m60	61t/m70	71t/m80	81t/m95

Uit tabel 6 van aselecte getallen nemen wij 10 getallen $\leqq 95$ (00 stelt 100 voor). Wij kiezen willekeurig rij 31 en lezen de eerste 10 getallen van 2 cijfers in horizontale richting af, waarbij wij geen herhalingen toestaan (steekproef zonder teruglegging).

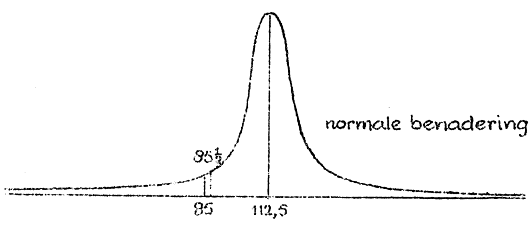
De getallen zijn:

59, 58, ~~00~~, 64, 78, 75, 56, ~~97~~, 88, ~~00~~, ~~88~~, 83, 55, 44.

De kinderen, die de steekproef vormen zijn afkomstig uit de volgende klassen:

nummer	klas	nummer	klas
44	2	64	4
55	3	75	5
56	3	78	5
58	3	83	6
59	3	88	6

4.



Wij toetsen de hypothese $H_0: p=\frac{1}{2}$ linkseenzijdig door de overschrijdingskans van de waarde

$$u = \frac{85\,\frac{1}{2} - 112{,}5}{\sqrt{225\,\frac{1}{2}\,\frac{1}{2}}} = -3{,}6$$

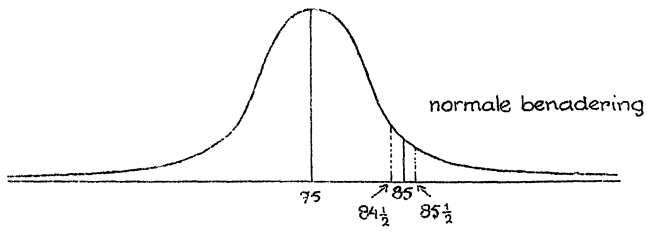
in tabel 1 op te zoeken:

$$k_\ell < 10^{-4}.$$

De tweezijdige overschrijdingskans is 2 x deze waarde:

$$k < 2 \times 10^{-4}.$$

De hypothese $p = \frac{1}{2}$ kan dus met grote stelligheid verworpen worden.



Vervolgens toetsen wij de hypothese $H_0: p = \frac{1}{3}$. De binomiale verdeling met parameters $p=\frac{1}{3}$ en $n=225$ wordt benaderd door een normale verdeling met parameters $\mu = \frac{1}{3}\cdot 225 = 75$ en $\sigma = \sqrt{225\,\frac{1}{3}\,\frac{2}{3}} = 7{,}07$. De tweezijdige overschrijdingskans is 2 x de kleinste éénzijdige overschrijdingskans van de waarde 85. De uitkomst (85) ligt rechts van $\mu = 75$, dus de rechtséénzijdige overschrijdingskans is de kleinste van de twee. Deze laatste wordt gevonden door de overschrijdingskans van

$$u = \frac{84\,\frac{1}{2} - 75}{7{,}07} = 1{,}34$$

in tabel 1 op te zoeken. Dit geeft $k_r = 0{,}09$, dus $k = 0{,}18$. De hypothese H_0: $p = \frac{1}{3}$ kunnen wij op grond hiervan niet verwerpen.

Daar $k_\ell > 0{,}5$ leidt linkséénzijdige toetsing zeker niet tot verwerping.

5. Wij toetsen de hypothese H_0 : In de ene stad komen niet systematisch meer ongevallen met dodelijke afloop voor dan in de andere stad. Indien deze hypothese juist is, is de kans, voor ieder jaar apart, om in A meer ongelukken te vinden dan in B gelijk aan de kans om in B meer ongelukken te vinden dan in A.

Wij kunnen deze hypothese dus toetsen met behulp van de tekentoets, waarbij de toetsingsgrootheid $\underline{x}$ gelijk is aan het aantal negatieve verschillen tussen het aantal ongelukken in A en B.

Wij vinden $x=3$: alleen in de jaren 1922, 1931 en 1932 zijn er in A minder ongelukken geweest dan in B.

De kritieke waarde van de tweezijdige tekentoets met $n=17$ en $\alpha=0{,}05$ bedraagt 4 (zie tabel 3). Het resultaat ligt dus in de kritieke zone. Wij verwerpen H_0, behoudens een onbetrouwbaarheid 0,05, ten gunste van de hypothese dat in A meer ongelukken gebeuren dan in B.

Aan het gevonden verschijnsel kunnen oorzaken van verschillende aard ten grondslag liggen. Stad A kan b.v. groter zijn dan B of ook de verkeersintensiteit kan er groter zijn, de wegen kunnen er slechter zijn, enz.

Daar hierover geen gegevens beschikbaar zijn, valt er geen gespecificeerde conclusie te trekken.

6. a) De fabrikant beweert, dat door behendigheid de kans op het verkrijgen van een prijs <u>groter</u> is dan 1/50.

Indien nu uit zijn resultaat, 8 prijzen bij 75 maal spelen, blijkt dat de ondergrens van het betrouwbaarheidsinterval voor de kans op winst groter is dan 1/50, dan zullen wij, behoudens een onbetrouwbaarheid 0,05, de fabrikant gelijk moeten geven.

De ondergrens lezen wij af uit tabel 2 en deze bedraagt voor $x=8$ $\quad n=75$ $\quad \alpha=0{,}05$ (eenzijdig) $\quad \theta_* = \frac{4{,}0}{75} = 0{,}053$.
Deze ondergrens is groter dan 0,02.

b) Vervolgens gaan wij na, hoeveel prijzen de fabrikant minstens gehaald zou moeten hebben, opdat hij, behoudens een onbetrouwbaarheid 0,05 zijn bewering nog juist kan staven.

Dan moet dus

$$\theta_* > 0{,}02$$

$$75\,\theta_* > 1{,}5 .$$

Voor $\alpha=0{,}05$ vinden wij dat de kleinste waarde van $n\theta_*$, die aan deze voorwaarde voldoet, gelijk is aan 2,0 en behoort bij $x=5$. De fabrikant moet dus minstens 5 prijzen behalen.

7. Voor de berekening van de verdeling van $\underline{x}$ onder de voorwaarde $\underline{x} \leqq 3$ maken wij gebruik van:

$$P[\underline{x}=x \mid \underline{x} \leqq 3] = \frac{P[\underline{x}=x \wedge \underline{x} \leqq 3]}{P[\underline{x} \leqq 3]} .$$

Nu is:

$$P[\underline{x}=0] = (\tfrac{2}{3})^5$$

$$P[\underline{x}=1] = 5\cdot\tfrac{1}{3}\cdot(\tfrac{2}{3})^4$$

$$P[\underline{x}=2] = 10(\tfrac{1}{3})^2(\tfrac{2}{3})^3$$

$$P[\underline{x}=3] = 10(\tfrac{1}{3})^3(\tfrac{2}{3})^2$$

$$P[\underline{x}=4] = 5(\tfrac{1}{3})^4\,\tfrac{2}{3}$$

$$P[\underline{x}=5] = (\tfrac{1}{3})^5$$

$$P[\underline{x}\leqq 3] = 1-\{P[\underline{x}=4 \text{ of } \underline{x}=5]\} = 1-\{5\cdot\tfrac{2}{3^5}+\tfrac{1}{3^5}\} = \tfrac{232}{3^5},$$

$$P[\underline{x}=0\,|\,\underline{x}\leqq 3] = \frac{P[\underline{x}=0 \wedge \underline{x}\leqq 3]}{P[\underline{x}\leqq 3]} = \frac{P[\underline{x}=0]}{P[\underline{x}\leqq 3]} = \frac{(2/3)^5}{\frac{232}{3^5}} = \frac{32}{232},$$

$$P[\underline{x}=1\,|\,\underline{x}\leqq 3] = \frac{5\cdot\frac{2^4}{3^5}}{\frac{232}{3^5}} = \frac{80}{232},$$

$$P[\underline{x}=2\,|\,\underline{x}\leqq 3] = \frac{80}{232},$$

$$P[\underline{x}=3\,|\,\underline{x}\leqq 3] = \frac{40}{232},$$

$$\mu_1 = \sum_{i=1}^{3} x_i p_i =$$

$$= 0\,\frac{32}{232} + 1\,\frac{80}{232} + 2\,\frac{80}{232} + 3\,\frac{40}{232} = 1{,}55$$

dus $\mu\{\underline{x}\,|\,\underline{x}\leqq 3\} = 1{,}55$

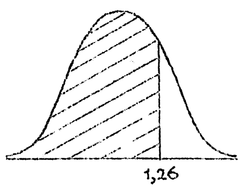
$$\mu_2 = \sum_{i=1}^{3} x_i^2 p_i =$$

$$= 0\,\frac{32}{232} + 1^2\,\frac{80}{232} + 2^2\,\frac{80}{232} + 3^2\,\frac{40}{232} = 3{,}28$$

$$\sigma^2 = \mu_2 - \mu^2 = 3{,}28 - 2{,}40 = 0{,}88$$

$$\sigma\{\underline{x}\,|\,\underline{x}\leqq 3\} = 0{,}94.$$

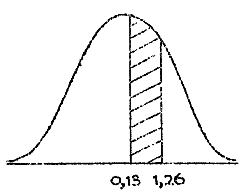
8. a) Kansen van de vorm $P[\underline{u}\geqq a] = P[\underline{u}>a]$ zijn direct uit de tabel der $N(0,1)$-verdeling af te lezen (tabel 1). Daar $\underline{u}$ continu verdeeld is, is immers voor iedere a: $P[\underline{u}=a]=0$.



$$P[\underline{u}\leqq 1{,}26] = 1 - P[\underline{u}\geqq 1{,}26]$$

$$= 1 - 0{,}1038$$

$$= 0{,}8962.$$



$$P[0{,}13<\underline{u}\leqq 1{,}26] = P[\underline{u}\geqq 0{,}13] - P[\underline{u}\geqq 1{,}26]$$

$$= 0{,}4483 - 0{,}1038$$

$$= 0{,}3445$$

$$P[\underline{u}\geqq 0{,}9] = 0{,}1841.$$

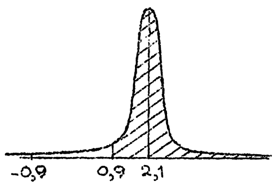
b) Gegeven: $\underline{x}$ is $N(\mu=2{,}1, \sigma=0{,}7)$-verdeeld. Bepaal $P[|\underline{x}| \geqq 0{,}9]$.

$$P[|\underline{x}| \geqq 0{,}9] = P[\underline{x} \geqq 0{,}9] + P[\underline{x} \leqq -0{,}9] =$$

$$= P\left[\frac{\underline{x}-2{,}1}{0{,}7} \geqq \frac{0{,}9-2{,}1}{0{,}7}\right] + P\left[\frac{\underline{x}-2{,}1}{0{,}7} \leqq \frac{-0{,}9-2{,}1}{0{,}7}\right] =$$

$$= P[\underline{u} \geqq -1{,}71] + P[\underline{u} \leqq -4{,}29].$$

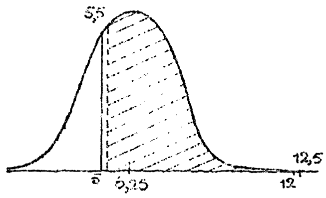
Hierin is de 2° term te verwaarlozen. Dus



$$P[\underline{u} \geqq -1{,}71] = 1 - P[\underline{u} \geqq 1{,}71] =$$

$$= 1 - 0{,}0436 = 0{,}9564.$$

c) De binomiale verdeling met $n=25$ en $p=0{,}25$ wordt benaderd door een normale verdeling met $\mu = np = 6{,}25$ en $\sigma = \sqrt{npq} = 2{,}17$.



Nu is dus bij benadering (let op de toepassing van de continuïteits-correctie):

$$P[5<\underline{y} \leqq 12] = P[\underline{y}>5] - P[\underline{y}>12] =$$

$$= P\left[\frac{\underline{y}-6{,}25}{2{,}17} > \frac{5+\frac{1}{2}-6{,}25}{2{,}17}\right] - P\left[\frac{\underline{y}-6{,}25}{2{,}17} > \frac{12+\frac{1}{2}-6{,}25}{2{,}17}\right] =$$

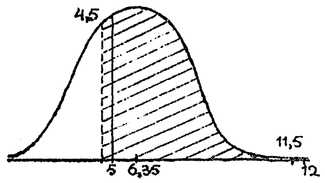
$$= P[\underline{u} > -0{,}35] - P[\underline{u} > 2{,}88] =$$

$$= 1 - \{P[\underline{u} > 0{,}35] + P[\underline{u} > 2{,}88]\} =$$

$$= 1 - \{0{,}3632 + 0{,}0020\} =$$

$$= 0{,}6348.$$

Analoog is:



$$P[5 \leqq \underline{y} < 12] = P[\underline{y} \geqq 5] - P[\underline{y} \geqq 12] =$$

$$= P\left[\frac{\underline{y}-6{,}25}{2{,}17} \geqq \frac{5-\frac{1}{2}-6{,}25}{2{,}17}\right] - P\left[\frac{\underline{y}-6{,}25}{2{,}17} \geqq \frac{12-\frac{1}{2}-6{,}25}{2{,}17}\right] =$$

$$= P[\underline{u} \geqq -0{,}81] - P[\underline{u} \geqq 2{,}42] =$$

$$= 1 - \{P[\underline{u} \geqq 0{,}81] + P[\underline{u} \geqq 2{,}42]\} =$$

$$= 1 - (0{,}2090 + 0{,}0078) =$$

$$= 0{,}7832.$$

9. a) De wh-verdeling van $\underline{x}$ is

$$P[\underline{x} = ia] = \frac{1}{k} \qquad i = 1, \ldots, k$$

$$\mathcal{E}\underline{x} = \mu = \sum_{i=1}^{k} ia\frac{1}{k} = \frac{a}{k}\sum_{i=1}^{k} i = \frac{a}{k}\,\frac{1}{2}k(k+1) = \frac{1}{2}(a+ka).$$

(Dus: Indien een stochastische grootheid een aantal aequidistante waarden met gelijke whn aanneemt ligt het gemiddelde midden tussen de 2 uiterste waarden.)

Nu is $\sigma^2\{\underline{x}\} = \mu_2 - \mu^2$

$$\mu_2 = \sum_{i=1}^{k} (ia)^2\frac{1}{k} = \frac{a^2}{k}\sum_{i=1}^{k} i^2 = \frac{a^2(2k+1)(k+1)}{6} \quad {}^{1)}$$

$$\sigma^2\{\underline{x}\} = \frac{a^2(2k+1)(k+1)}{6} - \frac{(k+1)^2a^2}{4} = \frac{a^2(k^2-1)}{12}.$$

---

1) Zie de opmerking verderop.

---

b) $\frac{1}{\underline{x}}$ neemt de waarden $\frac{1}{a}, \frac{1}{2a}, \ldots, \frac{1}{ka}$ aan, met gelijke whn.

$$P\left[\frac{1}{\underline{x}} = \frac{1}{ia}\right] = \frac{1}{k} \qquad i = 1, \ldots, k.$$

Voor $k = 5$ is:

$$\mu = \mathcal{E}\frac{1}{\underline{x}} = \sum_{i=1}^{5} \frac{1}{ia}\,\frac{1}{5} = \frac{1}{5a}\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}\right) = \frac{0{,}456}{a}$$

$$\mu_2 = \mathcal{E}\left(\frac{1}{\underline{x}}\right)^2 = \sum_{i=1}^{5} \frac{1}{i^2a^2}\,\frac{1}{5} = \frac{1}{5a^2}\left(1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25}\right) = \frac{0{,}292}{a^2}$$

$$\sigma^2\left\{\frac{1}{\underline{x}}\right\} = \mu_2 - \mu^2 = \frac{0{,}292}{a^2} - \frac{(0{,}456)^2}{a^2} = \frac{1}{2a^2}.$$

Opmerking.

$$\sum_{i=1}^{n} i^2 = 1+2+3+4+\cdots\cdots\cdots+n+ = \tfrac{1}{2}n(n+1)$$
$$+2+3+4+\cdots\cdots\cdots+n+ = \tfrac{1}{2}(n-1)(n+2)$$
$$+3+4+\cdots\cdots\cdots+n+ = \tfrac{1}{2}(n-2)(n+3)$$
$$\vdots$$
$$n = \tfrac{1}{2}\{n-(n-1)\}(n+n)$$
$$+$$
$$= \tfrac{1}{2}\sum_{i=1}^{n}(n-i+1)(n+i).$$

Dus

$$\sum i^2 = \tfrac{1}{2}\sum(n-i+1)(n+i)$$
$$2\sum i^2 = \sum(n^2-i^2+n+i) = \sum n^2 - \sum i^2 + \sum n + \sum i$$
$$3\sum i^2 = n^3+n^2+\tfrac{1}{2}n(n+1)$$
$$\sum i^2 = \frac{2n^3+3n^2+n}{6} = \frac{n(n+1)(2n+1)}{6}.$$

10. Met de toets van Wilcoxon toetsen wij de hypothese H_0: beide weefsels zijn even sterk. Wij voeren de toets tweezijdig uit omdat zowel A als B het sterkst kan zijn.

Voor de berekening van de U van Wilcoxon kunnen wij als volgt te werk gaan. Wij rangschikken de B-waarden in opklimmende volgorde van grootte en vervolgens noteren wij de A-waarden in de regels tussen die B-waarden, waartussen die A-waarden liggen.

De A-waarden behoeven dan niet van te voren naar opklimmende grootte te worden gerangschikt. (zie schema op blz. 30)

Wij tellen voor iedere A het aantal B-waarden die kleiner zijn dan die A. Zo zijn voor A = 681 de waarden B = 652, 654, 663, 676 kleiner. De bijdrage tot U bedraagt dus 4 (zie kolom 1).

De bijdrage tot U van iedere regel staat in kolom 4 genoteerd en deze wordt verkregen door het aantal A waarden in kolom 3 met het getal in kolom 1, dat op dezelfde regel staat, te vermenigvuldigen.

Zo krijgen wij:

$$U = 38$$
$$\mu = \tfrac{1}{2}mn = 50$$
$$\sigma^2\{U\} = \tfrac{1}{12}mn(m+n+1) = 175$$
$$\sigma\{U\} = 13,23.$$

De correctie voor gelijken behoeft bij de berekening van $\sigma^2\{\underline{U}\}$ in verband met het kleine aantal gelijken (3) niet te worden toegepast. De tweezijdige overschrijdingskans kan gevonden worden door in tabel 1 der N(0,1)-verdeling de overschrijdingskans op te zoeken van:

$$u = \frac{|U-\mu|-\frac{1}{2}}{\sigma\{\underline{U}\}} = \frac{|38-50|-\frac{1}{2}}{13,23} = 0,87.$$

Dit geeft: $k = 0,38$.

Conclusie: De hypothese H_0: beide weefsels zijn even sterk kan op grond van het resultaat van de proef, niet verworpen worden.

	B	A	Bijdrage tot U
0		629, 646	0
½	652		
1			
1½	654	654, 654	3
2		657	2
2½	663		
3			
3½	676		
4		681	4
4½	688		
5		697	5
5½	712		
6		713	6
6½	717		
7			
7½	728		
8			
8½	731		
9		736, 764	18
9½	783		
10			

totaal: U = 38

11. a) Twee gebeurtenissen A en B zijn stochastisch onafhankelijk indien

$$P[A \text{ en } B] = P[A].P[B].$$

---

b) De kans op gebeurtenis A: Het getrokken nummer is $\leqq 40$, is gelijk aan:

$$P[A] = \frac{40}{100}$$

en de kans op B: Het getrokken nummer is $\geqq 20$, is

$$P[B] = \frac{81}{100}.$$

Voor gebeurtenis A en B: Het getrokken nummer is

$$\leqq 40 \text{ en } \geqq 20,$$

geldt

$$P[A \text{ en } B] = \frac{21}{100}.$$

Nu is $P[A \text{ en } B] \neq P[A].P[B]$
want $\frac{21}{100} \neq \frac{40}{100} \cdot \frac{81}{100}$.
en A en B zijn dus niet stochastisch onafhankelijk.

---

c) De kans op A: het getrokken nummer $\leqq 40$, is

$$P[A] = \frac{40}{100}.$$

De kans op C: het getrokken nummer $\geqq n$, is:

$$P[C] = \frac{100-n+1}{100}.$$

Voor de gebeurtenis A en C: het getrokken nummer $\leqq 40$ en $\geqq n$ geldt, als $n \leqq 40$,

$$P[A \text{ en } C] = \frac{40-n+1}{100},$$

terwijl deze kans gelijk aan 0 is voor $n > 40$. Nu zijn de gebeurtenissen A en C positief afhankelijk indien:

$$P[A \text{ en } C] > P[A].P[C]$$

dus als $$\frac{40-n+1}{100} > \frac{40}{100} \cdot \frac{100-n+1}{100},$$

of wel $$n < 1.$$

Daar alleen de waarden 1,.......,100 aangenomen kunnen worden, zien wij, dat positieve afhankelijkheid niet op kan treden.

De gebeurtenissen A en C zijn negatief afhankelijk indien:

$$P[A \text{ en } C] < P[A].P[C]$$

dus als

$$\frac{40-n+1}{100} < \frac{40}{100}\,\frac{100-n+1}{100}$$

$$n > 1.$$

Voor $n > 40$ is $P[A \text{en} C]=0$, dus dan zijn A en C ook negatief afhankelijk. Voor $n=1$ is C zeker en dan zijn A en C stochastisch onafhankelijk.

12. De vispopulatie bestaat na de eerste vangst uit 50 gemerkte en $N-50$ ongemerkte vissen. De fractie gemerkte vissen is dus

$$\theta = \frac{50}{N}.$$

Met behulp van de tweede vangst ($n=100$, $x=9$) bepalen wij nu een tweezijdig betrouwbaarheidsinterval voor θ met onbetrouwbaarheid 0,05. Daartoe gebruiken wij tabel 2, want $x << n$ en wij gebruiken de kolommen met $\alpha = 0{,}025$ omdat de onbetrouwbaarheid in deze tabel éénzijdig staat opgegeven. Dit geeft:

$$4{,}0 < n\theta < 17{,}1$$

dus

$$\frac{4{,}0}{100} = 0{,}04 < \theta < \frac{17{,}1}{100} = 0{,}171,$$

dus

$$0{,}04 < \frac{50}{N} < 0{,}171,$$

$$294 < N < 1250,$$

behoudens een onbetrouwbaarheid 0,05.

13. a) $\underline{x}$ neemt de waarden $1, 2, \ldots, 100$ aan met kans $\frac{1}{100}$. Voor de berekening van $\sigma^2\{\underline{x}\}$ gebruiken wij de formule $\sigma^2 = \mu_2 - \mu^2$.

$$\mu = \sum x_i p_i = \frac{1+100}{2} = 50{,}5$$

$$\mu_2 = \sum x_i^2 p_i = \frac{1}{100} \sum x_i^2 = \frac{1}{100}\,\frac{100.101.201}{6} = 3383{,}5 \text{ (vgl.opgave 9)}$$

$$\sigma^2 = \mu_2 - \mu^2 = 3383{,}5 - 2550{,}25 = 833{,}25$$

$$\sigma = 28{,}866.$$

b) Wij beschouwen 3 achtereenvolgende waarnemingen. De mogelijkheden zijn, als wij het optreden van een waarde $\underline{x} \geq 80$ een succes noemen:

1. geen succes, dan is $\underline{y} > 3$
2. 1 succes, " " $\underline{y} \leq 3$
3. 2 successen, " " $\underline{y} \leq 2$
4. 3 successen, " " $\underline{y} = 1$.

Bij 3 waarnemingen moet één van deze 4 mogelijkheden optreden.

Nu is $P[\underline{y} > 3] = (0{,}79)^3$

dus $P[\underline{y} \leq 3] = 1 - (0{,}79)^3 = 0{,}507$.

14. De hypothese H_o, welke wij wensen te toetsen, luidt: De verkoop van gasfornuizen is, afgezien van toevallige fluctuaties, in beide winkels even groot.

Wij toetsen deze hypothese met de tweezijdige tekentoets, toegepast op de verschillen voor ieder der jaren. Dit geschiedt met de tekentoets en niet met de toets van Wilcoxon, omdat over de periode waarover de gegevens beschikbaar zijn, een verloop in de verkoop van gasfornuizen niet uitgesloten is en tweezijdig, omdat zowel A als B de grootste omzet kan hebben.

Wij vinden:

A	B	teken van A-B
209	163	+
210	204	+
182	183	-
200	238	-
197	203	-
232	273	-
232	274	-
253	264	-
277	227	+
227	285	-
284	244	+
222	235	-
293	290	+

De toetsingsgrootheid $\underline{x}$ is het aantal plustekens, dus $x = 5$
$n = 13$, $H_o : p = \frac{1}{2}$.
De tweezijdige overschrijdingskans vinden wij door in tabel 1 de overschrijdingskans op te zoeken van:

$$u = \frac{|5 - 6{,}5| - \frac{1}{2}}{\frac{1}{2}\sqrt{13}} = 0{,}56$$

$$\therefore K = 2 \times 0{,}2877 = 0{,}58$$

F. Direct uit tabel tekentoets te lezen.

Conclusie. Op grond van deze overschrijdingskans kunnen wij de getoetste hypothese niet verwerpen.

15. De kans op A "het getrokken nummer $\leqq a$" is gelijk aan

$$P[A] = \frac{a}{N}.$$

De kans op B "het getrokken nummer $\geqq b$" is gelijk aan

$$P[B] = \frac{N-b+1}{N}.$$

De gebeurtenissen A en B kunnen alleen dan tegelijk optreden indien $b < a$. Is dit niet zo, dan zijn zij, daar zij elkaar uitsluiten, stochastisch afhankelijk. Is $b < a$, dan is

$$P[A \text{ en } B] = \frac{a-b+1}{N}.$$

De gebeurtenissen A en B zijn nu stochastisch onafhankelijk indien:

$$\frac{a-b+1}{N} = \frac{a}{N} \cdot \frac{N-b+1}{N}.$$

Hieraan is alleen voldaan als $a = N$ of $b = 1$, d.w.z. als ofwel A zeker ofwel B zeker is. (vgl. opgave 11)

16. Noem het aantal paren gelijken $\underline{x}$; dan kan $\underline{x}$ dus de waarden $0, 1, 2, \ldots, 10$ aannemen. Wij berekenen $P[\underline{x} \geqq 1]$ als

$$1 - P[\underline{x} = 0].$$

Persoon A heeft aselect 10 nummers getrokken uit 100. De kans dat het 1° nummer, dat B trekt, ongelijk is aan één van de 10 nummers van A, is gelijk aan $\frac{90}{100}$. De kans dat het 2° nummer, dat B trekt, ongelijk is aan één van de 10 nummers van A, is gelijk aan $\frac{89}{99}$. Deze redenering voortzettend vinden wij:

$$P[\underline{x} = 0] = \frac{90}{100} \cdot \frac{89}{99} \cdot \frac{88}{98} \cdot \frac{87}{97} \cdot \ldots \cdot \frac{81}{91} = 0{,}33.$$

De gevraagde kans is dus:

$$P[\underline{x} \geqq 1] = 1 - 0{,}33 = 0{,}67.$$

Deze opgave kan ook als volgt worden opgelost:

Het aantal 10-tallen, dat uit de getallen 1 t/m 100 aselect te trekken is, bedraagt $\binom{100}{10}$.

Als wij nu de kans willen berekenen, dat er geen gelijke getallen voorkomen in de 10-tallen van A en B, dan bedenken wij dat B, nadat A heeft getrokken, dus nog 90 getallen tot zijn beschikking heeft, waaruit hij er 10 aselect mag trekken.

Het aantal mogelijke 10-tallen van deze aard bedraagt $\binom{90}{10}$, zodat dus

$$P[\underline{x}=0] = \binom{90}{10} / \binom{100}{10} = 0{,}33.$$

17. a) Wij toetsen, met behulp van de toets van Wilcoxon, de hypothese dat er geen verschil is in levensduur tussen beide soorten gloeilampen.

Aangezien de waarnemingen reeds gerangschikt zijn, kunnen wij direct de A en B waarnemingen samen in een opklimmende rij rangschikken.

Wij krijgen dan:

A B A B A A B A B A A B A B B A B B A B B B.

U wordt berekend door bij iedere A te tellen hoeveel B's ervóór staan, dus:

$$U = 0+1+2+2+3+4+4+5+7+9 = 37.$$

Verder is

$$\mu = \tfrac{1}{2}mn = \tfrac{1}{2}\,10.12 = 60$$

en

$$\sigma^2 = \tfrac{1}{12}nm(m+n+1) = \tfrac{1}{12}.10.12.23 = 230,$$

$$\sigma = 15{,}2.$$

De tweezijdige overschrijdingskans kan gevonden worden, door in tabel 1 van de $N(0,1)$-verdeling de overschrijdingskans op te zoeken van

$$u = \frac{|37-60| - \frac{1}{2}}{15{,}2} = 1{,}48.$$

Dit geeft $k = 0{,}14$.

Conclusie

Op grond van de gevonden overschrijdingskans kunnen wij de hypothese, dat er geen systematisch verschil in levensduur tussen de beide soorten gloeilampen is, niet verwerpen.

---

b) Indien de proef bekort wordt tot maximaal 1000 uren, vinden wij de U van Wilcoxon op onderstaande wijze (vergelijk opgave 10).

	B	A	Bijdrage tot U	Gelijken
0		625	0	
½	630			
1		637	1	
1½	683			
2		710, 770	4	
	780			
3		820	3	
	830			
4		843, 856	8	
	889			
5		920	5	
	970			
6				
$6+\frac{6}{2}=9$	6x >1000	2x >1000	18	8
12	m = 12	n=10	U=39	

Dus $U = 39,$

$$\mu = \tfrac{1}{2}\, mn = 60,$$

$$\sigma^2 = \frac{mn}{12}\left\{\frac{(m+n)^3-(1^3.g_1+2^3.g_2+3^3.g_3+\dots)}{(m+n)(m+n-1)}\right\} =$$

$$= \frac{10.12}{12}\left\{\frac{22^3-(1.14+8^3.1)}{22.21}\right\} = 219{,}09 .$$

$$\therefore\ \sigma = 14{,}8 .$$

De tweezijdige overschrijdingskans k wordt gevonden uit:

$$u = \frac{|39-60|-\frac{1}{2}}{14{,}8} = 1{,}385,$$

$$k = 0{,}17.$$

De conclusie is dus gelijkluidend aan die onder a.

<u>Opmerking</u>

De continuïteitscorrectie ($-\frac{1}{2}$ in de teller van u) is hier toegepast op dezelfde wijze als voor het geval, dat er geen gelijke waarnemingen zijn. Deze correctie berust op het feit, dat in het laatstgenoemde geval de intervallen tussen op elkaar volgende mogelijke waarden van $\underline{u}$ steeds de lengte 1 bezitten. Dit is, als er gelijken zijn, niet altijd meer zo. In sommige gevallen (zie b.v. opgave 50) blijven deze intervallen wel aan elkaar gelijk, maar worden langer; verkrijgen zij alle de lengte a, dan neme men als continuïteitscorrectie $-\frac{1}{2}a$ in plaats van $-\frac{1}{2}$. In veel gevallen zullen de gelijken echter de lengte van de intervallen tussen op elkaar volgende waarden niet op dezelfde wijze beïnvloeden: enerzijds kunnen ook niet gehele waarden voor $\underline{u}$ gevonden worden, anderzijds worden sommige waarden soms onmogelijk of minder waarschijnlijk dan zij eerst waren. Kortom, de toestand wordt door het optreden van gelijke waarnemingen minder overzichtelijk. Daar echter de overschrijdingskans, berekend met continuïteitscorrectie, steeds iets groter is dan die zonder deze correctie, valt het aan te bevelen ook in zulke onoverzichtelijke gevallen de gewone correctie te gebruiken. Men gaat op die wijze voorzichtiger te werk dan wanneer men hem, wegens de onoverzichtelijkheid, maar weglaat.

18. Daar de kans op succes gelijk aan p is, geldt

$$P[\underline{n}=1]=p,$$
$$P[\underline{n}=2]=qp,$$
$$P[\underline{n}=n]=q^{n-1}p,$$

dus

$$\mathcal{E}\underline{n}=\sum_{n=1}^{\infty} nq^{n-1}p.$$

Substitueer hierin $n=m+1$.

$$\begin{aligned}\mathcal{E}\underline{n} &= \sum_{m+1=1}^{\infty}(m+1)q^{m+1-1}p = \\ &= \sum_{m=0}^{\infty} mq^{m-1}qp + \sum_{m=0}^{\infty} q^m p = \\ &= q\sum_{m=1}^{\infty} mq^{m-1}p + \sum_{m=0}^{\infty} q^m p,\end{aligned}$$

daar de eerste term (voor $m=0$) van de eerste som gelijk aan 0 is. Nu is $\sum_{m=1}^{\infty} mq^{m-1}=\mathcal{E}\underline{n}$ (zie boven) en $\sum_{m=0}^{\infty} q^m = \frac{1}{1-q}$ (meetkundige reeks), dus

$$\mathcal{E}\underline{n}=q\mathcal{E}\underline{n}+\frac{p}{1-q}, \quad (1-q)\mathcal{E}\underline{n}=1$$
$$\mathcal{E}\underline{n}=\frac{1}{1-q}=\frac{1}{p}.$$

Verder is

$$\mathcal{E}\underline{n}^2 = \sum_{n=1}^{\infty} n^2 q^{n-1} p.$$

Substitueer wederom $n = m+1$.

$$\begin{aligned}\mathcal{E}\underline{n}^2 &= \sum_{m+1=1}^{\infty} (m+1)^2 q^{m+1-1} p = \\ &= \sum_{m=0}^{\infty} m^2 q^{m+1-1} p + 2\sum_{m=0}^{\infty} m q^{m+1-1} p + \sum_{m=0}^{\infty} q^m p = \\ &= q\sum_{m=1}^{\infty} m^2 q^{m-1} p + 2q\sum_{m=1}^{\infty} m q^{m-1} p + 1 = \\ &= q\,\mathcal{E}\underline{n}^2 + 2q\,\mathcal{E}\underline{n} + 1.\end{aligned}$$

Dus

$$(1-q)\,\mathcal{E}\underline{n}^2 = \frac{2q}{p} + 1,$$

$$\mathcal{E}\underline{n}^2 = \frac{2q+p}{p^2}.$$

Dus

$$\sigma^2\{\underline{n}\} = \mathcal{E}\underline{n}^2 - (\mathcal{E}\underline{n})^2 = \frac{2q+p}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}.$$

19. Indien uit de steekproef blijkt, dat de ondergrens van het betrouwbaarheidsinterval van het percentage ondeugdelijke exemplaren kleiner is dan 1%, dan is er geen reden tot reclameren. De ondergrens vinden wij in tabel 2 voor betrouwbaarheidsintervallen voor kleine kansen.
Voor $x = 5$ en $\alpha = 0{,}05$ is $\theta_* = \frac{2{,}0}{150} \times 100\% = 1{,}3\%$.

Hieruit zal men concluderen dat het percentage ondeugdelijke exemplaren in de partij goederen, behoudens een onbetrouwbaarheid 0,05 groter is dan 1,3% en men zal dus reclameren.

Men zal niet reclameren indien:

$$\theta_* < 0{,}01$$

dus

$$150\,\theta_* < 1{,}5.$$

Uit tabel 2 blijkt dat $x = 4$ de grootste waarde is, waarvoor hieraan nog voldaan is (dan is nl. $n\theta_* = 1{,}4$ voor $\alpha = 0{,}05$, terwijl voor $x = 5$ de ondergrens $n\theta_* = 2{,}0$ is, dus $> 1{,}5$).

Maximaal mogen dus 4 ondeugdelijke exemplaren in de steekproef van 150 exemplaren gevonden worden.

20. a) Toetsen wij de hypothese H_o dan is het onderscheidingsvermogen de kans op verwerpen van H_o indien H juist is.

$$\beta(H) \overset{def}{=} P[\underset{\sim}{x} \in Z \mid H].$$

($\in$ = behoort tot).

b) Aangezien de kritieke waarde $x=5$ de grootste waarde is die $\underset{\sim}{x}$ aan kan nemen als $n=5$, kunnen wij dus $P[\underset{\sim}{x} \in Z|H]$ berekenen als $P[\underset{\sim}{x}=5 \mid H]$. Wij kiezen voor H b.v. de volgende waarden van p: 0; 0,2; 0,3; 0,6; 0,7; 0,9 en 1.

$P[\underset{\sim}{x}=5 \mid 0] = 0$

$P[\underset{\sim}{x}=5 \mid 0{,}2] = 1.(0{,}2)^5 = 32.10^{-5}$

$P[\underset{\sim}{x}=5 \mid 0{,}3] = 1.(0{,}3)^5 = 243.10^{-5}$

$P[\underset{\sim}{x}=5 \mid 0{,}5] = 1.(0{,}5)^5 = 3125.10^{-5}$ (dit is de onbetrouwbaarheid van de toets).

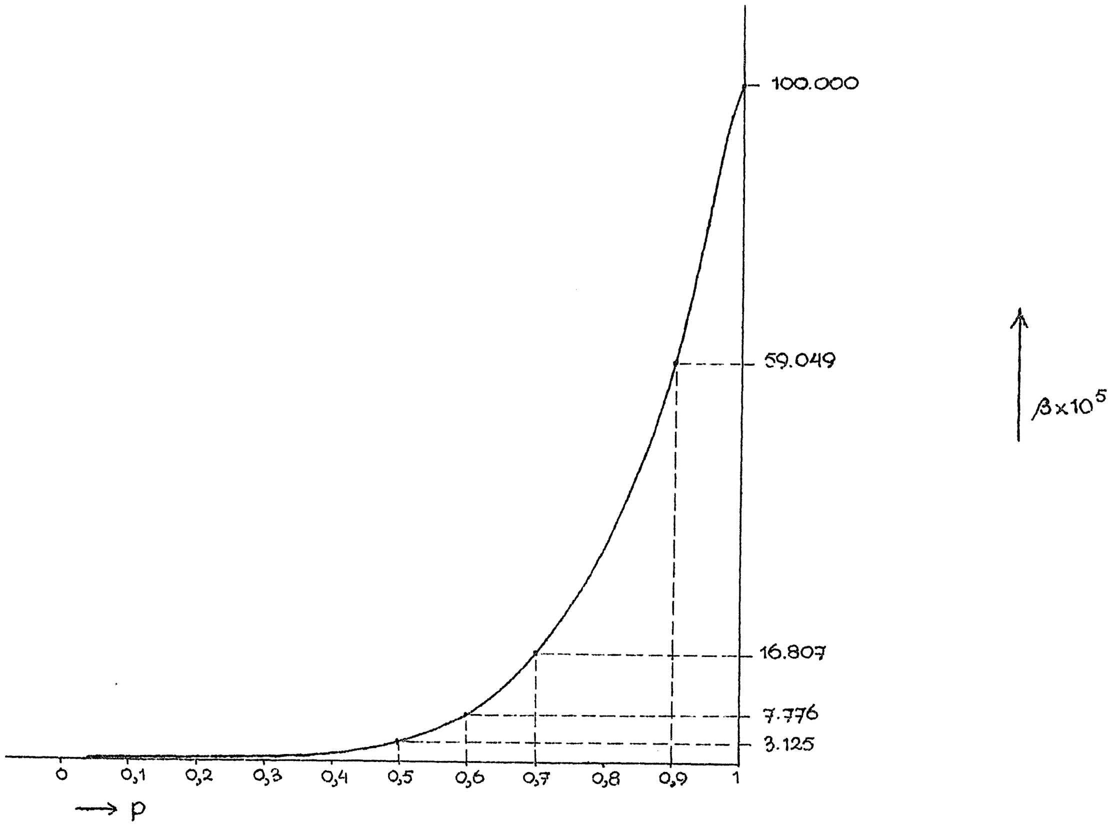
$P[\underset{\sim}{x}=5 \mid 0{,}6] = 1.(0{,}6)^5 = 7776.10^{-5}$

$P[\underset{\sim}{x}=5 \mid 0{,}7] = 1.(0{,}7)^5 = 16807.10^{-5}$

$P[\underset{\sim}{x}=5 \mid 0{,}9] = 1.(0{,}9)^5 = 59049.10^{-5}$

$P[\underset{\sim}{x}=5 \mid 1] = 1.$

Van dit onderscheidingsvermogen is onderstaande schets te maken.



Opmerking:

Indien wij het onderscheidingsvermogen willen benaderen, dan kunnen wij dit doen door voor iedere gekozen p te berekenen:

$$u = \frac{5-5p-\frac{1}{2}}{\sqrt{5pq}} \qquad \text{met } q = 1-p$$

en vervolgens in tabel 1 de eenzijdige overschrijdingskans van deze waarden van u op te zoeken.

De benadering is voor grote en kleine waarden van p niet erg goed; voor $p = 0{,}9$ b.v. vinden wij:

$$u = \frac{5-5\times 0{,}9-\frac{1}{2}}{\sqrt{5\times 0{,}9\times 0{,}1}} = 0.$$

De eenzijdige overschrijdingskans van deze waarde van u is 0,5, en dit is de benadering van het onderscheidingsvermogen. Het exact berekende onderscheidingsvermogen bedraagt in dit geval 0,59.

21. a) Wordt er voor een voorwerp geloot in welke groep het komt, dan is de kans om in A of B te komen voor beide $\frac{1}{2}$. Wij moeten dit dus alleen nog bewijzen voor de voorwerpen, waarom niet meer geloot wordt, omdat één van beide groepen reeds "vol" is. Een dergelijk voorwerp komt in B als A reeds vol is en in A als B reeds vol is.

Er blijft nu dus te bewijzen, dat de kans dat A het eerst vol is gelijk is aan de kans dat dit met B het geval is.

Bij iedere rij trekkingen, waarbij A het eerst vol is, kunnen wij door het verwisselen van A en B een dergelijke rij opschrijven waarbij B het eerst vol is. nl.

A A	B B
A B A	B A B
B A A	A B B.

Twee dergelijke rijen zijn steeds even lang en hebben dus, daar de munt zuiver is, gelijke wh. Voor de 2 in de eerste regel is deze b.v. 1/4 en voor de 4 andere 1/8. Derhalve is inderdaad de kans, dat A het eerst vol is, gelijk aan die voor B.

b) De verdeling is niet aselect. Immers de voorwerpen v_1 en v_2 vormen alleen dan groep A als wij A A verkrijgen en de kans daarop is 1/4. De voorwerpen v_1 en v_3 vormen tezamen groep A, als wij A B A krijgen en de kans daarop is 1/8. Deze kansen zijn niet gelijk.

22. a) De mogelijke tweetallen voor groep A zijn:

v_1v_2 , v_1v_3 , v_1v_4 , v_2v_3 , v_2v_4 en v_3v_4.

Deze 6 tweetallen moeten alle een kans 1/6 hebben om één groep te vormen. Wij moeten dus beschikken over een lotingsmechanisme met kans 1/6, en dit kunnen wij op de volgende wijze bereiken: Van de 8 mogelijke uitkomsten (met kans 1/8), die men bij 3 worpen met een zuivere munt kan krijgen, laten wij er 2 buiten beschouwing en aan ieder der overige voegen wij van te voren een combinatie van 2 voorwerpen toe. Bijv.:

mogelijke uitkomsten	tweetal	kans
K K M	v_1v_2	1/8
K M K	v_1v_3	1/8
M K M	v_1v_4	1/8
M M K	v_2v_3	1/8
K M M	v_2v_4	1/8
M K K	v_3v_4	1/8
K K K	} buiten beschouwing	1/8
M M M		1/8

Wij voeren nu zolang drietallen worpen uit tot één der 6 bovenste uitkomsten verkregen wordt.

Een dergelijke uitkomst geeft dan een verdeling van de groep van 4 voorwerpen in 2 groepen van 2 aan. Deze verdeling is aselect, want alle mogelijkheden hebben gelijke wh.

---

b) De berekening van $\mathcal{E}\underline{n}$ waarbij $\underline{n}$ gelijk is aan het aantal drietallen worpen, dat wij moeten verrichten om de aselecte verdeling tot stand te brengen, kan op analoge wijze als in opgave 18 geschieden.

Hier is: p = kans op succes = 3/4

q = kans op mislukking = 1/4.

$\underline{n}$ = aantal drietallen worpen, tot voor het eerst een succes optreedt.

Dus

$$\mathcal{E}\underline{n} = \frac{1}{p} = \frac{4}{3}.$$

De verwachting van het aantal worpen is dus $3 \times \frac{4}{3} = 4$.

23. a) Wij voeren de volgende notatie in

K = kruis

M = munt.

De munt is zuiver d.w.z. $P[M] = P[K] = \frac{1}{2}$. Derhalve is (vgl. opgave 18)

$$P[\underline{n}=n] = (\tfrac{1}{2})^{n-1}\tfrac{1}{2} = 2^{-n}.$$

De verwachting en de spreiding van de verdeling van $\underline{n}$ kunnen op dezelfde manier berekend worden als in opgave 18.
Wij krijgen dus: $\mathcal{E}\underline{n} = \frac{1}{p} = 2$

$$\text{en} \quad \sigma^2\{\underline{n}\} = \frac{q}{p^2} = 2.$$

---

b) $P[\underline{n}=\underline{m}]$ wordt als volgt berekend.
Voor speler A: $P[\underline{n}=i] = 2^{-i}$.
Voor speler B: $P[\underline{m}=i] = 2^{-i}$.
Dus: $P[\underline{n}=\underline{m}=i] = 2^{-2i}$

$$P[\underline{n}=\underline{m}] = \sum_{i=1}^{\infty} 2^{-2i} = \sum_{i=1}^{\infty} 4^{-i} = \frac{1}{3}.$$

---

c) In b) is al berekend dat

$$P[\underline{n}=\underline{m}] = \frac{1}{3}.$$

Dus is $P[\text{A wint of B wint}] = 1 - \frac{1}{3} = \frac{2}{3}$. Verder is de kans op winst voor A gelijk aan die voor B, dus is

$$P[\text{A wint}] = P[\text{B wint}] = \frac{1}{3}.$$

24. De linkeroverschrijdingskans van de waarde x is gelijk aan:

$$k_l = P[\underline{x} \leqq x].$$

De rechteroverschrijdingskans:

$$k_r = P[\underline{x} \geqq x].$$

De som van beide overschrijdingskansen:

$$\begin{aligned} k_l + k_r &= P[\underline{x} \leqq x] + P[\underline{x} \geqq x] = \\ &= P[\underline{x} < x] + P[\underline{x} = x] + P[\underline{x} > x] + P[\underline{x} = x] \\ &= 1 + P[\underline{x} = x]. \end{aligned}$$

Bij een continu verdeelde grootheid $\underline{x}$ is $P[\underline{x}=x] = 0$, zodat daarvoor $k_l + k_r = 1$ is.

25. Voor $P[\underline{x}^2 = x^2]$ kunnen wij schrijven

$$P[\underline{x}^2 = x^2] = P[\underline{x} = x \text{ of } \underline{x} = -x] = P[\underline{x} = -x] + P[\underline{x} = x].$$

Verder is

$$\begin{aligned} P[\underline{y} = y \text{ en } \underline{x}^2 = x^2] &= P[(\underline{y} = y \text{ en } \underline{x} = x) \text{ of } (\underline{y} = y \text{ en } \underline{x} = -x)] = \\ &= P[\underline{y} = y \text{ en } \underline{x} = x] + P[\underline{y} = y \text{ en } \underline{x} = -x] = \\ &= P[\underline{y} = y] . P[\underline{x} = x] + P[\underline{y} = y] . P[\underline{x} = -x] = \\ &= P[\underline{y} = y] \{ P[\underline{x} = x] + P[\underline{x} = -x] \} = \\ &= P[\underline{y} = y] \, P[\underline{x}^2 = x^2]. \end{aligned}$$

Volgens de definitie van stochastische onafhankelijkheid zijn $\underline{x}^2$ en $\underline{y}$ dus onafhankelijk.

26. a) Van de 4 mogelijke gevallen MK, KM, KK en MM van 2 worpen met een onzuivere munt beschouwen wij alleen de gevallen MK en KM.

Indien $P[K] = p$ en $P[M] = q = 1-p$,

dan is $P[KM] = P[MK] = P[K] . P[M] = pq$,

en $P[KM \text{ of } MK] = pq + qp = 2pq$,

$$P[KM \mid KM \text{ of } MK] = P[MK \mid KM \text{ of } MK] = \frac{pq}{2pq} = \frac{1}{2}.$$

Op deze wijze kan men dus een onzuivere munt "zuiver maken".

---

b) Werp telkens 2 x met de onzuivere munt en laat de resultaten KK en MM buiten beschouwing; indien MK voorkomt moet een voorwerp in groep A en indien KM voorkomt dan in groep B. Zo is de kans om in groep A te komen voor ieder voorwerp $\frac{1}{2}$ Bevat een der beide groepen A of B 3 voorwerpen, dan gaat de rest van de 6 voorwerpen in de andere groep (vgl. opgave 21 en 22).

27. Noem het aantal onjuiste conclusies $\underline{x}$ dan heeft $\underline{x}$ een binomiale verdeling met $n = 5$ en $p = \alpha$.

a) De kans op 5 goede conclusies is:

$$P[\underline{x} = 0] = (1-\alpha)^5$$

dus: $P[\underline{x} > 0] = 1-(1-\alpha)^5.$

---

b) $$P[\underline{x}=5]=\alpha^5.$$

c) Vul in $\alpha=0{,}02$ dan komt er

$$P[\underline{x}>0] = 1-(1-0{,}02)^5$$
$$= 0{,}096$$
$$P[\underline{x}=5] = (0{,}02)^5$$
$$= 3\times 10^{-9}.$$

d) Gevraagd wordt α zo te bepalen dat

$$P[\underline{x}=0] = 0{,}95$$

dus

$$(1-\alpha)^5 = 0{,}95$$
$$1-\alpha = \sqrt[5]{0{,}95} \approx 0{,}99$$
$$\alpha \approx 0{,}01 .$$

28. a) Indien de fabrikant een steekproef van 100 exemplaren neemt, moet hij op grond van het gevonden aantal defecte exemplaren kunnen beslissen dat de fractie defecte exemplaren in de partij niet groter is dan $\frac{3}{100}$, behoudens een onbetrouwbaarheid 0,1. De bovengrens van het betrouwbaarheidsinterval voor deze fractie mag dus niet groter zijn dan $\frac{3}{100}$. In tabel 2 voor betrouwbaarheidsintervallen voor kleine kansen vinden wij, als wij in de steekproef van 100 exemplaren geen enkel defect exemplaar hebben gevonden, voor de bovengrens:

$$\theta^* = \frac{2{,}3}{100} \qquad \text{behoudens} \quad \alpha=0{,}1 .$$

Deze bovengrens is kleiner dan $\frac{3}{100}$. Vinden wij in de steekproef van 100 exemplaren 1 defect exemplaar, dan geldt voor de bovengrens

$$\theta^* = \frac{3{,}9}{100} \qquad \text{behoudens} \quad \alpha=0{,}1 .$$

Deze bovengrens is groter dan $\frac{3}{100}$. De fabrikant zal dus die partijen goedkeuren, waarbij hij geen enkel defect exemplaar in de steekproef van 100 exemplaren heeft gevonden. Bij 1 of meer defecte exemplaren in de steekproef van 100 exemplaren wordt de partij afgekeurd. De kans dat hij partijen met 3% of meer defecte exemplaren ten onrechte goedkeurt is dan $\leq 0{,}1$.

b) De steekproefgrootte stellen wij n. De fabrikant wil dus, indien hij 1 defect exemplaar in de steekproef vindt, de partij nog net goedkeuren, d.w.z. concluderen dat

$$\theta^* \leqq \frac{3}{100}.$$

Dus (zie tabel 2 voor $x = 1$ en $\alpha = 0,1$)

$$\frac{3,9}{n} \leqq \frac{3}{100}$$

$$n \geqq 130.$$

De steekproef moet dus minstens 130 exemplaren bevatten.

29. Indien n en x beide groot zijn kunnen wij de onder- en bovengrens van het betrouwbaarheidsinterval voor θ berekenen uit de formules 4' en 5' (zie formules)

$$\theta^*_*(x) = \frac{1}{n}\left\{ x \pm c'\sqrt{\frac{x(n-x)}{n}} \right\},$$

waarin c' een constante is, die afhangt van de onbetrouwbaarheidsdrempel. Voor $\alpha = 0,01$ (tweezijdig) is $c' = 2,58$ (zie tabel B). De lengte van het betrouwbaarheidsinterval voor θ is nu:

$$\theta^* - \theta_* = \frac{2c'}{n}\sqrt{\frac{x(n-x)}{n}} = \frac{2.2,58}{n}\sqrt{\frac{x(n-x)}{n}}.$$

Bij gelijkblijvende θ en α zal het betrouwbaarheidsinterval nauwer worden, indien n toeneemt. Als $\theta = 0,4$ dan zal $\frac{x}{n}$ in de buurt van 0,4 vallen, dus zal x ongeveer $0,4\,n$ zijn. Willen wij het betrouwbaarheidsinterval dan een lengte 0,1 geven, dan vinden wij de grootte van de steekproef uit bovenstaande formule:

$$0,1 \approx \frac{2.2,58}{n}\sqrt{\frac{0,4\,n(n-0,4n)}{n}},$$

$$n \approx \frac{2^2\ 2,58^2 .\ 0,24}{0,1^2}$$

$$n \approx 639.$$

30. a) Van de partij goederen van 10.000 exemplaren zijn 500 ondeugdelijk. De kans dat een aselect uit de partij gekozen exemplaar ondeugdelijk is, is dus

$$\frac{500}{10.000} = 0,05.$$

---

b) Voor het nabootsen van het nemen van een steekproef van 40 exemplaren uit deze partij, moeten wij dus beschikken over een lotingsmechanisme met kans 0,05 op "een ondeugdelijk exemplaar". Van de getallen 0 tot 100 kiezen wij er 5, die de ondeugdelijke exemplaren voorstellen. Wij kiezen b.v. 00, 01, 02, 03 en en 04. Vervolgens nemen wij 40 getallen, zonder herhalingen uit tabel 6 van de aselecte getallen (kolommen 11 en 12) en tellen daarin het aantal getallen dat < 05 is.

32	86	~~84~~	75	65
51	69	01 ←	23	38
47	93	~~87~~	94	~~65~~
20	68	~~47~~	18	35
66	62	95	13	~~07~~
78	~~93~~	03 ←	19	82
81	11	07	84	98
~~84~~	44	06	54	
61	17	99	42	
00 ←	87	43	14	

In de aldus genomen steekproef zijn de aangestreepte exemplaren ondeugdelijk. Dit zijn er 3.

Uit tabel 2 voor betrouwbaarheidsintervallen voor kleine kansen vinden wij voor het tweezijdig betrouwbaarheidsinterval met $\alpha = 0,05$

$$n\theta_* = 0,6 \qquad n\theta^* = 8,8$$

dus $\theta_* = \frac{0,6}{40} = 0,015$ en $\theta^* = \frac{8,8}{40} = 0,22$.

Dit interval bevat inderdaad de werkelijke kans 0,05 voor het trekken van een ondeugdelijk exemplaar.

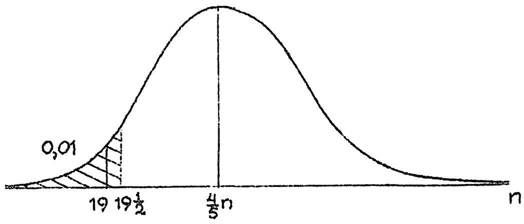
Opmerking. Het niet toelaten van herhalingen in de rij van 40 aselecte getallen komt overeen met het nemen van een steekproef van 40 exemplaren zonder teruglegging, zoals dit in de practijk gewoonlijk gedaan wordt. Dit betekent, strict genomen, dat de kans op het trekken van een ondeugdelijk exemplaar afhankelijk is van de resultaten der reeds eerder getrokken exemplaren. Deze kans is dus niet constant, maar daar de omvang van de steekproef (40) klein is in vergelijking met die van de gehele partij (10.000) is dit effect zo klein, dat het zonder bezwaar verwaarloosd kan worden.

31. Stel de organisator neemt n vuurpijlen mee naar het vuurwerk, hij heeft dus $n-20$ reserve pijlen.

Iedere pijl heeft een kans $p = 4/5$ om een succes te zijn. Noem het aantal successen $\underline{x}$, dan moet de kans, dat 19 of minder pijlen goed zijn $\leqq 0,01$ zijn; dus

$$P[\underline{x} \leqq 19] \leqq 0,01.$$

$\underline{x}$ bezit een binomiale verdeling met parameters n en $p = 4/5$. Wij benaderen deze met een normale verdeling met parameters $\mu = 4/5\,n$ en $\sigma = \sqrt{n\frac{4}{5}\cdot\frac{1}{5}} = \frac{2}{5}\sqrt{n}$.



Dus

$$P\left[\underline{u} \leq \frac{19\frac{1}{2} - \frac{4}{5}n}{\frac{2}{5}\sqrt{n}}\right] \leq 0{,}01.$$

Uit tabel 1 van de $N(0,1)$-verdeling volgt, dat het linkerlid gelijk is aan 0,01 als

$$-2{,}33 = \frac{19\frac{1}{2} - \frac{4}{5}n}{\frac{2}{5}\sqrt{n}}.$$

$\therefore\ n_1 = 30{,}80$ en $n_2 = 19{,}27$, waarvan alleen n_1 in aanmerking komt. Dus 31 - 20 = 11 reservepijlen.

32. a) Indien uit het steekproefresultaat blijkt dat de partij goederen meer dan 2% defecte exemplaren bevat, dan zal de fabrikant de partij afkeuren. Ligt de ondergrens van het betrouwbaarheidsinterval voor θ, de kans op een defect exemplaar, hoger dan 2%, dan zal behoudens een onbetrouwbaarheid $\alpha = 0{,}05$ het percentage defecte exemplaren in de partij goederen groter dan 2% zijn. De ondergrens van het betrouwbaarheidsinterval vinden wij uit tabel 2 voor $x=5$, $n=150$ en $\alpha = 0{,}05$ (eenzijdig): $n\theta_* = 2{,}0$ dus $\theta_* = 1{,}33\%$.

Het vermoeden van de fabrikant is dus juist, want hij zal nu niet tot afkeuren van de partij overgaan.

---

b) De fabrikant zal de partij afkeuren indien hij uit de steekproef van 150 exemplaren vindt dat:

$$\theta_* > 0{,}02$$

dus als

$$150\,\theta_* > 0{,}02 . 150 = 3.$$

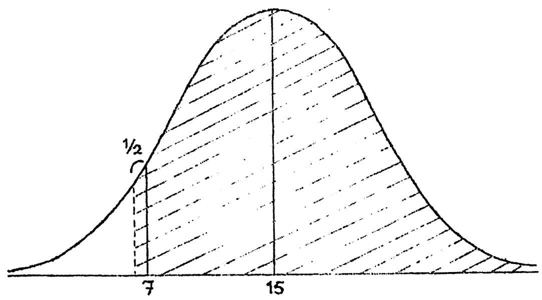
Bij opzoeken in tabel 2 vinden wij voor $x=6$ de waarde $n\theta_* = 2{,}6$, die < 3 is, terwijl voor $x=7$ de waarde 3,3 voor $150\,\theta_*$ gevonden wordt.

Indien de fabrikant dus 7 of meer defecte exemplaren in een steekproef van 150 exemplaren vindt, zal hij de partij afkeuren, waarbij hij dan een kans $\leqq 0,05$ heeft om een partij die $\leqq 2\%$ defecten bevat, ten onrechte af te keuren.

---

c) Bij 7 of meer defecten in de steekproef van 150 exemplaren wordt de partij afgekeurd.

De kans om de partij af te keuren, als het werkelijke percentage defecten 10 bedraagt, is dus gelijk aan de rechteroverschrijdingskans van 7 (gearceerde deel in de figuur).



Voor de berekening gebruiken wij de normale benadering waarin

$$\mu = np = 150\ 0{,}1 = 15,$$

$$\sigma = \sqrt{npq} = \sqrt{150 . 0{,}1 . 0{,}9} = 3{,}68 .$$

$$u = \frac{7-15-\frac{1}{2}}{3{,}68} = \frac{-8{,}5}{3{,}68} = -2{,}31 .$$

In tabel 1 van de N(0,1)-verdeling vinden wij nu

$$k_l = 0{,}01$$

en :

$$k_r = 0{,}99 .$$

M.a.w. de kans op afkeuren van een partij met minstens 10% defecten is, indien bij 7 of meer defecten in een steekproef van 150 exemplaren afgekeurd wordt, gelijk aan 0,99.

Hoewel wij hier met kleine kansen te maken hebben, gaan wij toch eens na, welke oplossing wij met behulp van de normale verdeling vinden.

a) Wij toetsen de hypothese $H_0: p=0{,}02$ tegen de alternatieve hypothesen $p>0{,}02$. Ligt het resultaat $x=5$ bij $n=150$, onder $H_0: p=0{,}02$ in de rechterkritieke zone met $\alpha=0{,}05$, dan zullen wij H_0 verwerpen ten gunste van $H: p>0{,}02$. De binomiale verdeling met $n=150$ en $p=0{,}02$ wordt nu benaderd door een normale verdeling met $\mu=np=3$ en $\sigma=\sqrt{npq}=1{,}71$.

$$u=\frac{5-3-\frac{1}{2}}{1{,}71}=0{,}88.$$

De rechter eenzijdige overschrijdingskans, opgezocht in tabel 1 der $N(0,1)$-verdeling bedraagt $k_\tau=0{,}19$.
Conclusie: Het steekproefresultaat leidt niet tot verwerping van $H_0: p=0{,}02$ en het vermoeden van de fabrikant is juist.

---

b) Hoeveel defecte exemplaren had de fabrikant moeten vinden opdat hij wel tot verwerping van $H_0: p=0{,}02$ zou zijn overgegaan? Stel dit aantal x, dan moet x dus onder H_0 in de rechterkritieke zone liggen, d.w.z.

$$P[\underline{x} \geqq x \mid p=0{,}02] \leqq 0{,}05$$

ofwel

$$P\left[\frac{\underline{x}-3-\frac{1}{2}}{1{,}71} \geqq \frac{x-3-\frac{1}{2}}{1{,}71} \,\middle|\, p=0{,}02\right] \leqq 0{,}05.$$

Uit tabel 1 vinden wij nu voor $\frac{x-3\frac{1}{2}}{1{,}71}$ de waarde 1,65.

$$\therefore\ x=6{,}32,$$

d.w.z., daar $\underline{x}$ alleen gehele waarden aanneemt, dat voor $\underline{x}\geqq 7$ tot verwerping wordt overgegaan.

In dit geval zijn de uitkomsten, verkregen volgens tabel 2 en volgens de normale benadering gelijk.

33. Bij gelijke winstkansen voor beide spelers moet de winstkans voor A $\frac{1}{2}$ zijn.

Noemen wij het aantal malen, dat A wint $\underline{x}$, dan kunnen wij de hypothese $H_0: p=\frac{1}{2}$ toetsen door de tweezijdige overschrijdingskans van $x=65$ berekenen, waarbij de binomiale verdeling met $p=\frac{1}{2}$ en $n=100$ benaderd wordt door een normale verdeling met $\mu=50$ en $\sigma=\sqrt{100\cdot\frac{1}{2}\cdot\frac{1}{2}}=5$. Wij toetsen in dit geval tweezijdig, omdat zowel A als B een kans $>\frac{1}{2}$ kunnen hebben om te winnen.

$$u=\frac{|65-50|-\frac{1}{2}}{5}=2{,}9.$$

In tabel 1 van de $N(0,1)$-verdeling vinden wij een tweezijdige overschrijdingskans $k=0{,}004$.

Conclusie: De hypothese $H_0: p=\frac{1}{2}$ wordt verworpen wegens de kleine waarde van k ten gunste van de hypothese, dat A een grotere winstkans heeft dan B.

---

b) Nu wordt gevraagd, de grens van het rechterdeel der tweezijdige kritieke zone van $\underline{x}$ voor $\alpha=0{,}05$ te berekenen. Dit geschiedt als volgt: Wij zoeken naar de kleinste $\underline{x}$ waarvoor

$$\frac{x-50-\frac{1}{2}}{5} \geqq \xi_{0,05}=1{,}96 \quad \text{dus:} \quad x \geqq 60{,}3.$$

Daar $\underline{x}$ alleen gehele waarden aan kan nemen, mag A dus hoogstens 60 maal winnen zonder dat wij bij een onbetrouwbaarheidsdrempel $\alpha=0{,}05$ de hypothese $H_0: p=\frac{1}{2}$ behoeven te verwerpen. Uit tabel 3 van de tekentoets is dit resultaat rechtstreeks af te lezen.

---

c) De kans op verwerping van $H_0: p=\frac{1}{2}$ wordt groot als p dicht bij 0 of dicht bij 1 ligt. Wij berekenen nu, hoe groot $p(>\frac{1}{2})$ moet zijn, om een kans 0,8 te verkrijgen op verwerping van H_0 ten gunste van de juiste hypothese $p>\frac{1}{2}$. Dit is (zie ook b)

$$\beta(p)=P[\underline{x}\geqq 61\,|\,p]=P\left[\frac{\underline{x}-100p}{10\sqrt{pq}} \geqq \frac{61-100p}{10\sqrt{pq}}\,\Big|\,p\right],$$

waarin nu $\frac{\underline{x}-100p}{10\sqrt{pq}}=\underline{u}$ bij benadering $N(0,1)$ verdeeld is. Voor die waarden van p, waarbij voldaan is aan:

$$P\left[\underline{u}\geqq \frac{61-100p}{10\sqrt{pq}}\right]\approx 0{,}8$$

is dus het onderscheidingsvermogen $\approx 0{,}8$. Nu is $P[\underline{u}\geqq -0{,}84]=0{,}8$ zoals uit tabel 1 blijkt. Dus is $\beta(p)\approx 0{,}8$ als

$$\frac{61-100p}{10\sqrt{pq}}=-0{,}84, \qquad 100{,}71\,p^2-122{,}71\,p+37{,}21=0,$$
$$p_1=0{,}65\,, \quad p_2=0{,}57\,.$$

Voor $p=0{,}65$ is het onderscheidingsvermogen $\approx 0{,}8$. Voor $p=0{,}57$ is het onderscheidingsvermogen $\approx 0{,}2$ want deze wortel van de vierkantsvergelijking is ingevoerd als $6{,}1-10p=+0{,}84\sqrt{pq}$ en $P[\underline{u}\geqq 0{,}84]=0{,}2$; uit symmetrieoverwegingen volgt dat ook voor $p\approx 1-0{,}65=0{,}35$ het onderscheidingsvermogen $\approx 0{,}8$ bedraagt.

Opmerking. Formeel verstaat men in de toetsingstheorie onder het "onderscheidingsvermogen" de totale kans op verwerping van H_0, zowel indien dit ten gunste van de juiste hypothese (b.v. $p>\frac{1}{2}$) plaats vindt als ten gunste van de verkeerde (b.v. $p<\frac{1}{2}$ als in werkelijkheid $p>\frac{1}{2}$). Dit is een onpractische definitie van het onderscheidingsvermogen, daar men bij toepassingen weinig gesteld zal zijn op het optellen van de kans op een falikant verkeerde conclusie en die op de juiste. Wij hebben daarom de laatste ongeveer gelijk aan 0,8 gemaakt. Daar de kans op een falikant verkeerde conclusie in de regel zeer gering is, is het verschil in de uitkomsten van deze methode en de gebruikelijke meestal slechts klein.

34. Algemeen geldt: Indien de gebeurtenis A optreedt kan de gebeurtenis B wel of niet optreden. Dit schrijven wij als:

$$A \equiv (A \text{ en } B) \text{ of } (A \text{ en niet } B).$$

Dus is volgens het vierde axioma:

$$P[A] = P[A \text{ en } B] + P[A \text{ en niet } B] - P[(A \text{ en } B) \text{ en } (A \text{ en niet } B)]$$

de gebeurtenissen $(A \text{ en } B)$ en $(A \text{ en niet } B)$ sluiten elkaar echter uit, dus $P[(A \text{ en } B) \text{ en } (A \text{ en niet } B)] = 0$.
Gegeven is nu dat als A optreedt, B ook optreedt, dus $P[A \text{ en niet } B] = 0$
dus $P[A] = P[A \text{ en } B]$.
Geheel analoog vinden wij

$$P[B] = P[A \text{ en } B],$$

zodat

$$P[A] = P[B].$$

35. a) Indien A optreedt kan B in het algemeen wel of niet optreden. Wij schrijven dit als:

$$A \equiv (A \text{ en } B) \text{ of } (A \text{ en niet } B).$$

Analoog aan opgave 34 vinden wij dus

$$P[A] = P[A \text{ en } B] + P[A \text{ en niet } B],$$

maar de laatste term is in dit geval gelijk aan 0, daar B steeds optreedt als A optreedt. Dus

$$P[A] = P[A \text{ en } B].$$

---

b) Nu uitgaande van

$$P[B] = P[B \text{ en } A] + P[B \text{ en niet } A]$$

vinden wij, daar wij reeds bewezen hebben dat,

$$P[A] = P[A \text{ en } B];$$

$$\underline{P[B] \geq P[A].}$$

c) $P[A \text{ of } B] = P[A] + P[B] - P[A \text{ en } B]$......(axioma 4)

Nu is:

$$P[A] = P[A \text{ en } B]$$

dus $P[A \text{ of } B] = P[B]$.

---

d) Te bewijzen is nu: $P[B|A] = 1$

$$P[B|A] \stackrel{\text{def}}{=} \frac{P[A \text{ en } B]}{P[A]}.$$

Nu is $P[A \text{ en } B] = P[A]$

dus $P[B|A] = \dfrac{P[A]}{P[A]} = 1.$

36. a) De gebeurtenissen A en B sluiten elkaar uit, dus $P[A \text{ en } B] = 0$. A en B zijn stochastisch onafhankelijk indien geldt:

$$P[A \text{ en } B] = P[A].P[B]$$

en daar $P[A \text{ en } B] = 0$ is hieraan dus alleen voldaan indien òf $P[A] = 0$ òf $P[B] = 0$.

---

b en c) Voor de gebeurtenis A schrijven wij weer:

$$A \equiv (A \text{ en } B) \text{ of } (A \text{ en niet } B)$$

dus daar de gebeurtenis (A en niet B) onmogelijk is, is

$$P[A] = P[A \text{ en } B]$$

dus als B altijd optreedt als A optreedt, zijn A en B niet stochastisch onafhankelijk, tenzij $P[B] = 1$.

37. a) Er zijn $\binom{n}{k}$ verschillende verdelingen in twee groepen A en B mogelijk. Nummeren wij al deze mogelijkheden van $1, \ldots, \binom{n}{k}$ en beschikken wij over een loterij met $\binom{n}{k}$ verschillende uitkomsten met gelijke whn, dan kunnen wij de aselecte verdeling dus tot stand brengen, door één maal uit deze loterij te trekken.

In principe kan men dit met een zuivere munt als volgt bereiken. Zij h het kleinste getal, waarvoor

$$2^h \geqq \binom{n}{k}$$

(b.v. $n=5$ en $k=2$ geeft $\binom{n}{k} = 10$ en $h=4$). Werpen wij nu h maal met een zuivere munt, dan zijn er 2^h verschillende worpreeksen mogelijk. B.v. voor $h=4$ de 16 worpreeksen:

```
K K K K
K K K M
K K M K
K M K K
M K K K
K K M M
K M K M
K M M K
M K K M
M K M K
- - - - - - - -
M M K K
   :
   :
M M M M .
```

Deze bezitten alle dezelfde wh en wij zoeken er nu willekeurig $\binom{n}{k}$ uit en laten de rest buiten beschouwing. In het voorbeeld nemen wij b.v. de eerste 10. Aan ieder hiervan voegen wij één der mogelijke verdelingen van de n voorwerpen in de twee groepen A en B toe. Vervolgens gaan wij inderdaad n maal werpen met een zuivere munt en dit experiment herhalen wij tot wij één der $\binom{n}{k}$ gekozen reeksen verkrijgen, en brengen de daaraan toegevoegde verdeling in de groepen A en B tot stand. Alle mogelijke verdelingen hebben dan dezelfde wh.

Deze methode eist voor grote n zeer veel tijd.

---

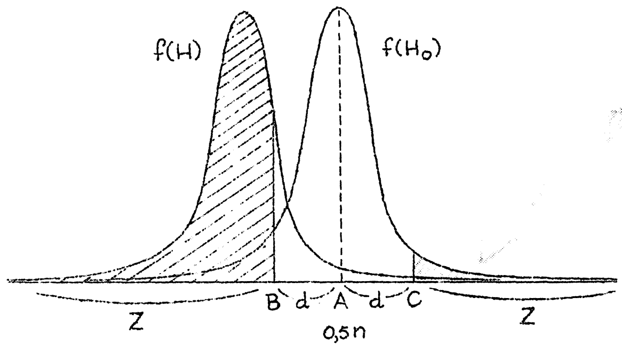
b) Een snellere methode is te verkrijgen door te werken met aselecte getallen 0,.....,9. Wij nummeren dan de n voorwerpen van $1, \ldots, n$ en nemen uit een lijst van aselecte getallen de eerste k getallen (zonder herhaling) die kleiner dan $n+1$ zijn, waarbij 00 als 100 gelezen wordt. Is b.v. $n=83$ en $k=17$, dan nemen wij paren aselecte getallen, slaan alle getallen > 83 over, evenals getallen, die wij reeds gehad hebben, en nemen de eerste 17 getallen ≤ 83.

Tabel 6 van aselecte getallen, rij 23 geeft het volgende:
~~97~~, 12, 54, 03, 48, ~~87~~, 08, 33, 14, 17, 21, 81, 53, ~~92~~, 50, 75, 23, 76, 20 en 47.

De doorgestreepte getallen zijn > 83 en komen niet in aanmerking. Herhalingen kwamen in deze rij niet voor. De gevonden getallen wijzen tezamen de voorwerpen van groep A aan.

Opmerking: De onder a) beschreven methode kan hier ook worden toegepast, maar is veel omslachtiger. De onder b) beschreven methode kan ook met behulp van een zuivere munt toegepast worden, indien men de nummers der voorwerpen in het tweetallige stelsel schrijft.

38. Wij berekenen eerst de kritieke waarden B en C van de tekentoets voor n experimenten en $\alpha = 0{,}05$, waarbij wij de normale benadering toepassen.



Nu wordt d (zie figuur) zo gekozen, dat

$$P[\underline{x} \in Z \mid H_0] = P[|\underline{x} - \tfrac{1}{2}n| > d \mid H_0] = \alpha$$

of ook, met $\alpha = 0{,}05$:

$$P\left[\left|\frac{\underline{x} - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}}\right| > \frac{d}{\frac{1}{2}\sqrt{n}}\right] = P\left[|\underline{u}| > \frac{d}{\frac{1}{2}\sqrt{n}}\right] = 0{,}05 .$$

De waarde van ξ_α, waarvoor

$$P[|\underline{u}| > \xi_\alpha] = 0{,}05$$

is $\xi_\alpha = 1{,}96$, zoals uit tabel 1 blijkt. Dus moet gelden

$$\frac{d}{\frac{1}{2}\sqrt{n}} = 1{,}96 ,$$
$$\therefore\ d = 0{,}98\sqrt{n} .$$

De kritieke waarden zijn dus

$$B = \tfrac{1}{2}n - 0{,}98\sqrt{n} \qquad \text{en} \qquad C = \tfrac{1}{2}n + 0{,}98\sqrt{n} .$$

Willen wij nu -behoudens een wh 0,1 -de afwijking van $p = 0{,}5$ ontdekken, dan moet dus het onderscheidingsvermogen van de tekentoets voor $p = 0{,}2$ gelijk aan 0,9 zijn. Het onderscheidingsvermogen, dat gedefinieerd is als:

$$\beta(H) \overset{\text{def}}{=} P[\underline{x} \in Z \mid H]$$

is in dit geval, waarin wij de binomiale verdeling nu benaderen door een normale verdeling met $\mu = 0{,}2n$ en $\sigma = 0{,}4\sqrt{n}$ gelijk aan:

$$\beta(0{,}2) = F_N\left(\frac{\frac{1}{2}n - 0{,}98\sqrt{n} - 0{,}2n}{0{,}4\sqrt{n}}\right) + \left\{1 - F_N\left(\frac{\frac{1}{2}n + 0{,}98\sqrt{n} - 0{,}2n}{0{,}4\sqrt{n}}\right)\right\},$$

waarin $F_N(x)$ de verdelingsfunctie van de normale (0,1)-verdeling voorstelt. Nu is (zie figuur) in dit geval

$$F_N\left(\frac{\frac{1}{2}n+0,98\sqrt{n}-0,2n}{0,4\sqrt{n}}\right)\approx 1$$

zodat wij voor het onderscheidingsvermogen bij benadering mogen schrijven:

$$\beta(0,2)\approx F_N\left(\frac{\frac{1}{2}n-0,98\sqrt{n}-0,2n}{0,4\sqrt{n}}\right) = 0,9$$

nu is $F_N(1,28) = 0,9$ zodat dus:

$$\frac{\frac{1}{2}n-0,98\sqrt{n}-0,2n}{0,4\sqrt{n}} \geqq 1,28,$$

$$\frac{0,3\sqrt{n}-0,98}{0,4} \geqq 1,28,$$

$$0,3\sqrt{n} \geqq 1,492,$$

$$n \geqq 24,7.$$

Conclusie: Wij moeten n minstens gelijk aan 25 nemen.

39. Laat $\underline{z}$ het aantal caramels zijn, dat het jongetje krijgt, $\underline{x}$ het aantal bonnetjes in de gekochte caramels, $\underline{y}'$ het aantal bonnetjes in de eerste portie gewonnen caramels en $\underline{y}''$ het aantal bonnetjes in de tweede portie gewonnen caramels.

a) Als het jongetje precies 5 caramels krijgt, vindt hij dus geen enkel bonnetje in zijn gekochte caramels dus dan is $x = 0$:

$$P[\underline{z}=5] = P[\underline{x}=0] = (0,8)^5 = 0,33.$$

De kans, dat hij meer dan 5 caramels krijgt is dus

$$P[\underline{z}>5] = 1 - P[\underline{z}=5] = 1-0,33 = 0,67.$$

---

b) Als het jongetje precies 6 caramels krijgt, heeft hij dus in de 5 gekochte caramels 1 bonnetje gevonden, terwijl de daarvoor cadeau gekregen caramel geen bonnetje bevatte. Dus $x=1$ en $y'=0$.

$$P[\underline{z}=6] = P[\underline{x}=1 \text{ en } \underline{y}'=0] = P[\underline{x}=1]\,P[\underline{y}'=0] =$$

$$\binom{5}{1}(0,2)^1(0,8)^4\cdot(0,8)^1 =$$

$$5\,(0,2)^1\cdot(0,8)^5 = 0,33.$$

---

c) Het jongetje kan op 2 manieren precies 7 caramels krijgen:

c 1. doordat hij in de 5 gekochte caramels 2 bonnetjes vindt, waarop hij 2 caramels krijgt, die geen van beide bonnetjes bevatten,

c 2. doordat hij in de 5 gekochte caramels 1 bonnetje vindt, waarop hij 1 caramel krijgt, die ook een bonnetje bevat, terwijl de zevende caramel geen bonnetje bevat.

c 1. $$P_1[\underline{z}=7] = P[\underline{x}=2 \text{ en } \underline{y}'=0] = P[\underline{x}=2] \,.\, P[\underline{y}'=0] =$$
$$= \binom{5}{2}(0{,}2)^2 (0{,}8)^3 (0{,}8)^2 =$$
$$= 10(0{,}2)^2 (0{,}8)^5.$$

c 2. $$P_2[\underline{z}=7] = P[\underline{x}=1 \text{ en } \underline{y}'=1 \text{ en } \underline{y}''=0] =$$
$$= P[\underline{x}=1] \,.\, P[\underline{y}'=1] \,.\, P[\underline{y}''=0] =$$
$$= \binom{5}{1}(0{,}2)^1 (0{,}8)^4 (0{,}2) . (0{,}8) =$$
$$= 5(0{,}2)^2 (0{,}8)^5.$$
$$P[\underline{z}=7] = P_1[\underline{z}=7] + P_2[\underline{z}=7] =$$
$$= 10(0{,}2)^2 (0{,}8)^5 + 5(0{,}2)^2 (0{,}8)^5 =$$
$$= 15(0{,}2)^2 (0{,}8)^5 = 0{,}20.$$

---

d) $$P[\underline{z}>7] = P[\underline{z}>5] - \{P[\underline{z}=6] + P[\underline{z}=7]\} =$$
$$= 0{,}67 - (0{,}33 + 0{,}20) = 0{,}14.$$

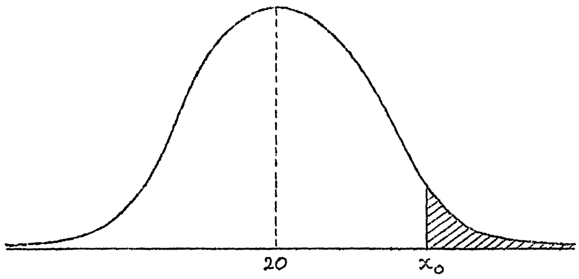
40. a) Noem $\underline{x}$ het aantal bonnetjes in de steekproef, dan is $\underline{x}$ binomiaal verdeeld, dus

$$\mathcal{E}\underline{x} = np = 100 \times 0{,}2 = 20$$

$$\sigma^2\{\underline{x}\} = npq = 100 \times 0{,}2 \times 0{,}8 = 16\,, \quad \sigma\{\underline{x}\} = 4.$$

---

b) De winkelier protesteert alleen als hij <u>meer</u> bonnetjes in de caramels vindt, dan redelijkerwijs met het opgegeven percentage in overeenstemming te brengen is. In dat geval moet hij te veel caramels cadeau geven, waardoor hij schade lijdt. Wij gebruiken hier dus de rechtséénzijdige kritieke zone, waarvan de kritieke waarde, x_0, gevraagd wordt met onbetrouwbaarheidsdrempel $\alpha = 0{,}05$;



$$P[\underline{x} \geqq x_o] = P[\underline{u} \geqq \frac{x_o-\frac{1}{2}-20}{4}] = 0{,}05.$$

Uit tabel 1 van de normale verdeling blijkt, dat hieraan voldaan is als

$$\frac{x_o-20-\frac{1}{2}}{4} = 1{,}65$$

$$\therefore\ x_o = 27{,}1$$

dus indien de winkelier 28 of meer caramels met bonnetjes in zijn steekproef vindt, zal hij gaan protesteren.

41. a) Laat n het aantal caramels zijn dat het jongetje in 3 weken gekocht heeft, en laat $\underline{x}$ het aantal prijzen zijn. Dan is:

$$P[\underline{x}=0] = (0{,}8)^n = 0{,}0687\ ;$$

$$\therefore\ n = 12.$$

Hij kon dus van één weekgeld 4 caramels kopen.

---

b) Laat n het aantal caramels zijn, dat de vader voor zijn zoontje koopt en laat daarvan $\underline{x}$ het aantal prijzen zijn, dan moet

$$P[\underline{x} \geqq 1] = 0{,}99,$$

dus

$$P[\underline{x}=0] = 0{,}01 = (0{,}8)^n\ ;$$

$$\therefore\ n = 20{,}6.$$

De vader koopt 21 caramels.

42. Algemeen geldt: Indien de gebeurtenis A optreedt kan de gebeurtenis B wel of niet optreden. Dit schrijven wij als:

$$A \equiv (A \text{ en } B) \text{ of } (A \text{ en } \bar{B}). \qquad (\bar{B} \equiv \text{niet } B).$$

Dus volgens het 4e axioma:

$$P[A] = P[A \text{ en } B] + P[A \text{ en } \bar{B}] - P[(A \text{ en } B) \text{ en } (A \text{ en } \bar{B})].$$

De gebeurtenissen (A en B) en (A en $\bar{B}$) sluiten elkaar echter uit, dus $P[(A \text{ en } B) \text{ en } (A \text{ en } \bar{B})] = 0$.
Gegeven is nu $P[A \text{ en } B] = P[A].P[B]$.
Vullen wij dit in de vorige vergelijking in, dan komt er

$$P[A] = P[A].P[B] + P[A \text{ en } \bar{B}]$$

$$P[A](1 - P[B]) = P[A \text{ en } \bar{B}]$$

maar:

$$1 - P[B] = P[\bar{B}],$$

dus:

$$P[A].P[\bar{B}] = P[A \text{ en } \bar{B}].$$

De gebeurtenissen A en $\bar{B}$ zijn dus ook stochastisch onafhankelijk.

43. a) Volgens de definitie van stochastische onafhankelijkheid zou $P[\underline{z} = z \mid \underline{x} = x]$ voor iedere waarde van x dezelfde waarde moeten bezitten. Nu is b.v. $P[\underline{z} = 2 \mid \underline{x} = 1] = 1/6$ en $P[\underline{z} = 2 \mid \underline{x} = 3] = 0$.
$\underline{x}$ en $\underline{z}$ zijn dus niet stochastisch onafhankelijk.

---

b) Wij hebben de volgende 4 mogelijkheden:

x	y	z
even	even	even
even	oneven	oneven
oneven	even	oneven
oneven	oneven	even

die ieder een kans 1/4 hebben.

$$\therefore \; P[\underline{z} = \text{oneven}] = 1/2.$$

Beschouwen wij nu alleen de gevallen waarbij $\underline{x}$ even is dan zien wij dat ook dan in de helft van de gevallen z oneven is, zodat dus

$$P[\underline{z}=\text{oneven} \mid \underline{x}=\text{even}] = 1/2$$

$$\therefore\ P[\underline{z}=\text{oneven}] = P[\underline{z}=\text{oneven} \mid \underline{x}=\text{even}].$$

De gebeurtenissen A en B zijn dus stochastisch onafhankelijk.

---

c) Een voorwaardelijke stochastische grootheid $\underline{w}$ onder de voorwaarde A geven wij aan met $(\underline{w} \mid A)$. Dan is:

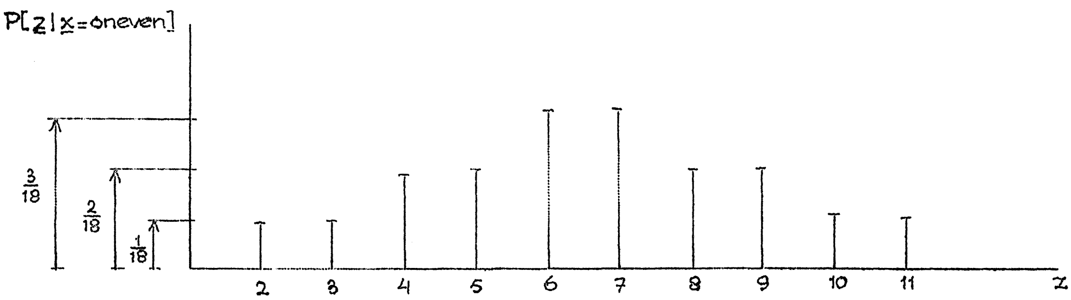
$$(\underline{z} \mid \underline{x}=\text{oneven}) = (\underline{x} \mid \underline{x}=\text{oneven}) + \underline{y}.$$

De voorwaardelijke waarschijnlijkheidsverdeling van $\underline{z}$, onder de voorwaarde dat $\underline{x}$ oneven is, kan als volgt worden berekend.

x	y	z
1	1	2
	2	3
	3	4
	4	5
	5	6
	6	7
3	1	4
	2	5
	3	6
	4	7
	5	8
	6	9
5	1	6
	2	7
	3	8
	4	9
	5	10
	6	11

Er zijn dus 18 mogelijkheden, ieder met gelijke kans 1/18

De waarschijnlijkheidsverdeling van $\underline{z}$ heeft de volgende vorm:



Uit de waarschijnlijkheidsverdeling van $(\underline{z} \mid \underline{x}=\text{oneven})$ berekenen wij het 1e en het 2e moment.

$$\mathcal{E}(\underline{z}\,|\,\underline{x}=\text{oneven}) = \sum z_i P[\underline{z}=z_i\,|\,\underline{x}=\text{oneven}] =$$

$$= (2+3+10+11)\tfrac{1}{18} + (4+5+8+9)\tfrac{2}{18} + (6+7)\tfrac{3}{18} =$$

$$= 6{,}5\ ,$$

$$\mathcal{E}(\underline{z}^2\,|\,\underline{x}=\text{oneven}) = \sum z_i^2 P[\underline{z}=z_i\,|\,\underline{x}=\text{oneven}] =$$

$$= (2^2+3^2+10^2+11^2)\tfrac{1}{18} + (4^2+5^2+8^2+9^2)\tfrac{2}{18} + (6^2+7^2)\tfrac{3}{18} =$$

$$= 47{,}83\ ;$$

$$\sigma^2\{\underline{z}\,|\,\underline{x}=\text{oneven}\} = \mu_2-\mu^2 = 5{,}58$$

$$\sigma\{\underline{z}\,|\,\underline{x}=\text{oneven}\} = 2{,}36.$$

Voor de onvoorwaardelijke verdeling van $\underline{z}$ geldt:

$$\mathcal{E}\underline{z} = \mathcal{E}\underline{x} + \mathcal{E}\underline{y} = 3\tfrac{1}{2} + 3\tfrac{1}{2} = 7$$

en

$$\sigma^2\{\underline{z}\} = \sigma^2\{\underline{x}\} + \sigma^2\{\underline{y}\} = \tfrac{35}{12} + \tfrac{35}{12} = 5{,}83$$

dus

$$\sigma\{\underline{z}\} = 2{,}41.$$

Het gemiddelde en de spreiding zijn dus voor de voorwaardelijke verdeling kleiner dan voor de onvoorwaardelijke verdeling van $\underline{z}$.

44. a) De formule voor de toets van Student is:

$$t_2 = \frac{\bar{x}-\bar{y}}{\sqrt{S_1^2+S_2^2}}\sqrt{\frac{nm(n+m-2)}{n+m}}\ ,$$

waarin

$$S_1^2 = \sum(x_i-\bar{x})^2 = \sum x_i^2 - 2\bar{x}\sum x_i + n\bar{x}^2 = \sum x_i^2 - n\bar{x}^2 =$$

$$= \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

en evenzo

$$S_2^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{m}$$

is.

	x	x^2	y	y^2
	0,87	0,7569	0,52	0,2704
	0,75	0,5625	1,94	3,7636
	0,03	0,0009	0,53	0,2809
	0,17	0,0289	1,02	1,0404
	0,67	0,4489	0,94	0,8836
	1,92	3,6864	0,06	0,0036
	0,23	0,0529	0,25	0,0625
	2,14	4,5796	2,05	4,2025
	2,69	7,2361	1,93	3,7249
	1,31	1,7161	1,17	1,3689
			1,38	1,9044
			1,43	2,0449
			1,74	3,0276
Σx_i	10,78	$\Sigma x_i^2 = 19{,}07$ $\quad \Sigma y_i$	14,96	$\Sigma y_i^2 = 22{,}58$
$(\Sigma x_i)^2$	116,21	$n = 10$ $\quad (\Sigma y_i)^2$	223,80	$m = 13$
$\Sigma x_i/n$	1,078	$\Sigma y_i/m$	1,151	
S_1^2	7,45	S_2^2	5,36	

$$\sqrt{S_1^2 + S_2^2} = 3{,}58.$$

Deze waarden in de formule voor de toets van Student ingevuld geven:

$$t_2 = \frac{1{,}078 - 1{,}151}{3{,}58} \sqrt{\frac{10.13.21}{23}} = 0{,}22.$$

Het aantal vrijheidsgraden is $m+n-2 = 21$. De tweezijdige overschrijdingskans, opgezocht in tabel 4 van de Student-verdeling bij $\nu = 21$ bedraagt: $0{,}8 < k < 0{,}9$.

<u>Conclusie</u>: De hypothese "er is geen verschil tussen de gemiddelden van de beide verdelingen" kan met de toets van Student niet verworpen worden.

---

b) Voor het tweezijdig betrouwbaarheidsinterval voor het verschil van de twee gemiddelden geldt de formule:

$$\bar{x} - \bar{y} - t_\alpha \sqrt{S_1^2 + S_2^2} \sqrt{\frac{n+m}{nm(n+m-2)}} < \mu_1 - \mu_2 < \bar{x} - \bar{y} + t_\alpha \sqrt{S_1^2 + S_2^2} \sqrt{\frac{n+m}{nm(n+m-2)}}.$$

De waarde van t_α, opgezocht in tabel 4, bedraagt voor $\nu = 21$ en $\alpha = 0{,}05$ (tweezijdig) $t_\alpha = 2{,}08$. Ingevuld in bovenstaande formule:

$$1{,}078 - 1{,}151 - 2{,}08 \times 3{,}58 \sqrt{\frac{23}{10.13.21}} < \mu_1 - \mu_2 < 1{,}078 - 1{,}151 + 2{,}08 \times 3{,}58 \sqrt{84 \times 10^{-4}}$$

$$-0{,}755 < \mu_1 - \mu_2 < +0{,}609.$$

Dit is het gezochte betrouwbaarheidsinterval.

---

c) Voor de berekening van de U van Wilcoxon kunnen wij te werk gaan als in opgave 10.

	x	y			Bijdrage tot U
0					
	0,03				
1		0,06			1
	0,17				
2					
	0,23				
3		0,52	0,53	0,25	9
	0,67				
4					
	0,75				
5					
	0,87				
6		1,02	0,94	1,17	18
	1,31				
7		1,38	1,43	1,74	21
	1,92				
8		1,94	2,05	1,93	24
	2,14				
9					
	2,69				
10					

n=10 m=13 U = 73

Wij vinden $U = 73$

$$\mathcal{E}\underline{u} = \tfrac{1}{2}mn = 65$$

$$\sigma^2\{\underline{u}\} = \tfrac{1}{12}mn(m+n+1) = 260 \qquad \sigma\{\underline{U}\} = 16{,}12$$

$$\frac{|u - \mathcal{E}\underline{u}| - \frac{1}{2}}{\sigma\{\underline{U}\}} = \frac{7{,}5}{16{,}12} = 0{,}47.$$

De tweezijdige overschrijdingskans, opgezocht in tabel 1, bedraagt $k = 0{,}64$.

Conclusie: De hypothese, dat de beide steekproeven uit eenzelfde verdeling komen, kan met de toets van Wilcoxon niet verworpen worden.

45. a) Het resultaat van de trekking van r ballen kan op de onderstaande wijze in een z.g. 2 x 2- tabel worden gerangschikt.

	getrokken	niet-getrokken	
wit	$\underline{a}$	$\underline{c}$	n
zwart	$\underline{b}$	$\underline{d}$	m
	r	s	N

hierin is

$$\underline{a} + \underline{c} = n$$

$$\underline{b} + \underline{d} = m$$

$$r + s = n + m = N.$$

De n witte ballen kunnen op $\binom{n}{a}$ verschillende wijzen in a getrokken en c niet-getrokken ballen verdeeld worden.

De m zwarte ballen kunnen op $\binom{m}{b}$ verschillende wijzen in b getrokken en d niet-getrokken ballen verdeeld worden.

Er zijn dus $\binom{n}{a}\binom{m}{b}$ verschillende wijzen om de N ballen in 2 groepen te verdelen, waarvan de ene bestaat uit a witte en b zwarte ballen, terwijl de andere het restant is. In totaal zijn er $\binom{N}{r}$ verschillende r-tallen te vormen uit N ballen. Hiervan geven er, volgens het bovenstaande $\binom{n}{a}\binom{m}{b}$ het gewenste resultaat (d.w.z. a witte ballen) en daar (wegens het aselecte karakter van de trekking) alle r-tallen gelijke kans bezitten om getrokken te worden is

$$P[\underline{a}=a] = \binom{n}{a}\binom{m}{b}\Big/\binom{N}{r}.$$

Opmerking:

Daar $a+b=r$ en $n+m=N$ en $\sum_a P[\underline{a}=a]=1$ als over alle mogelijke waarden van a gesommeerd wordt, is:

$$\sum_a \binom{n}{a}\binom{m}{r-a}\Big/\binom{n+m}{r} = 1.$$

Uit de 2 x 2- tabel zien wij dat a in het algemeen niet alle waarden van 0 tot n kan aannemen; indien $\tau > m$ is de kleinste waarde die a aan kan nemen gelijk aan $\tau - m$ en als $\tau < n$ is de grootste waarde van a gelijk aan τ.

---

b) $\mathcal{E}\underline{a} = \sum a p_a$ met $p_a = P[\underline{a} = a]$.

Dus met $b = \tau - a$

$$\mathcal{E}\underline{a} = \sum_a a \binom{n}{a}\binom{m}{b} \Big/ \binom{N}{\tau} =$$

$$= \sum_a \frac{a n (n-1)!}{a(a-1)!(n-a)!} \; \frac{\tau(\tau-1)!(N-\tau)!}{N(N-1)!} \binom{m}{b} =$$

$$= \frac{n\tau}{N} \sum_a \frac{(n-1)!}{(a-1)!(n-a)!} \; \frac{(\tau-1)!(N-\tau)!}{(N-1)!} \binom{m}{b} =$$

$$= \frac{n\tau}{N} \sum_a \binom{n-1}{a-1}\binom{m}{b} \Big/ \binom{N-1}{\tau-1} = \frac{n\tau}{N},$$

want volgens bovenstaande opmerking is de som in het 2e lid gelijk aan 1, daar $n-1+m = N-1$ en $a-1+b = \tau-1$ is.

Voor de berekening van $\mathcal{E}\underline{a}^2$ overwegen wij, dat

$$\mathcal{E}\underline{a}^2 = \mathcal{E}\,\underline{a}(\underline{a}-1) + \mathcal{E}\underline{a}.$$

Nu is (weer met $b = \tau - a$):

$$\mathcal{E}\underline{a}(\underline{a}-1) = \sum_a a(a-1) \frac{\binom{n}{a}\binom{m}{b}}{\binom{N}{\tau}} =$$

$$= \sum_a \frac{a(a-1)n!}{a!(n-a)!} \; \frac{\tau!(N-\tau)!}{N!} \binom{m}{b} =$$

$$= \sum_a \frac{a(a-1)n(n-1)(n-2)!}{a(a-1)(a-2)!(n-a)!} \; \frac{\tau(\tau-1)(\tau-2)!(N-\tau)!}{N(N-1)(N-\cancel{\tau}\,2)!} \binom{m}{b} =$$

fout

$$= \frac{n(n-1)\tau(\tau-1)}{N(N-1)} \sum_a \frac{\binom{n-2}{a-2}\binom{m}{b}}{\binom{N-2}{\tau-2}} =$$

$$= \frac{n\tau(n-1)(\tau-1)}{N(N-1)},$$

want $a-2+b=\tau-2$ en $n-2+m=N-2$, dus de som is weer gelijk aan 1, evenals boven.
Dus is:

$$\mathcal{E}\underline{a}^2 = \frac{n\tau(n-1)(\tau-1)}{N(N-1)} + \frac{\tau n}{N} =$$

$$= \frac{\tau n}{N(N-1)} \left\{ \tau(n-1)+m \right\}$$

en

$$\sigma^2 = \mu_2 - \mu^2 =$$

$$= \frac{\tau n\{\tau(n-1)+m\}}{N(N-1)} - \frac{\tau^2 n^2}{N^2} =$$

$$= \frac{nm\tau s}{N^2(N-1)} \ .$$

46. Wij denken de hele vaas leeggetrokken en de ballen in volgorde van trekking op een rij gelegd. Alle mogelijke rangschikkingen van de $n+m$ ballen in deze rij bezitten dan dezelfde wh. Het aantal verschillende wijzen, om n rode ballen op $n+m$ plaatsen te leggen is $A=\binom{n+m}{n}$. Ligt er op de $(k+1)$e plaats een witte bal, dan zijn er voor de rode ballen nog $n+m-1$ plaatsen over en daarover kunnen zij op $B=\binom{n+m-1}{n}$ wijzen verdeeld worden. Van de A mogelijkheden met gelijke wh, leiden er dus B tot het gewenste resultaat, zodat de gezochte kans gelijk is aan

$$\frac{B}{A} = \frac{\binom{n+m-1}{n}}{\binom{m+n}{n}} = \frac{m}{n+m} \ .$$

Opmerking:
Het is strikt genomen, niet nodig A en B expliciet op te schrijven en op elkaar te delen. Het is nl. voldoende op te merken, dat A en B beide onafhankelijk van k zijn, dus voor iedere k gelijk. Wij kunnen dus $k+1=1$ nemen en verkrijgen dan dezelfde uitkomst. De kans, dat de eerste bal een witte is, is echter $\frac{n}{n+m}$, zoals direct uit het aantal rode en witte ballen volgt.

47. $\underline{x}$ en $\underline{y}$ zijn beide binomiaal verdeeld.

(1) $$P[\underline{x}=x] = \binom{n}{x} p^x q^{n-x} \qquad \text{met} \qquad p+q=1$$

en

(2) $$P[\underline{y}=y] = \binom{m}{y} p_1^{y} q_1^{m-y} \qquad \text{met } p_1+q_1=1.$$

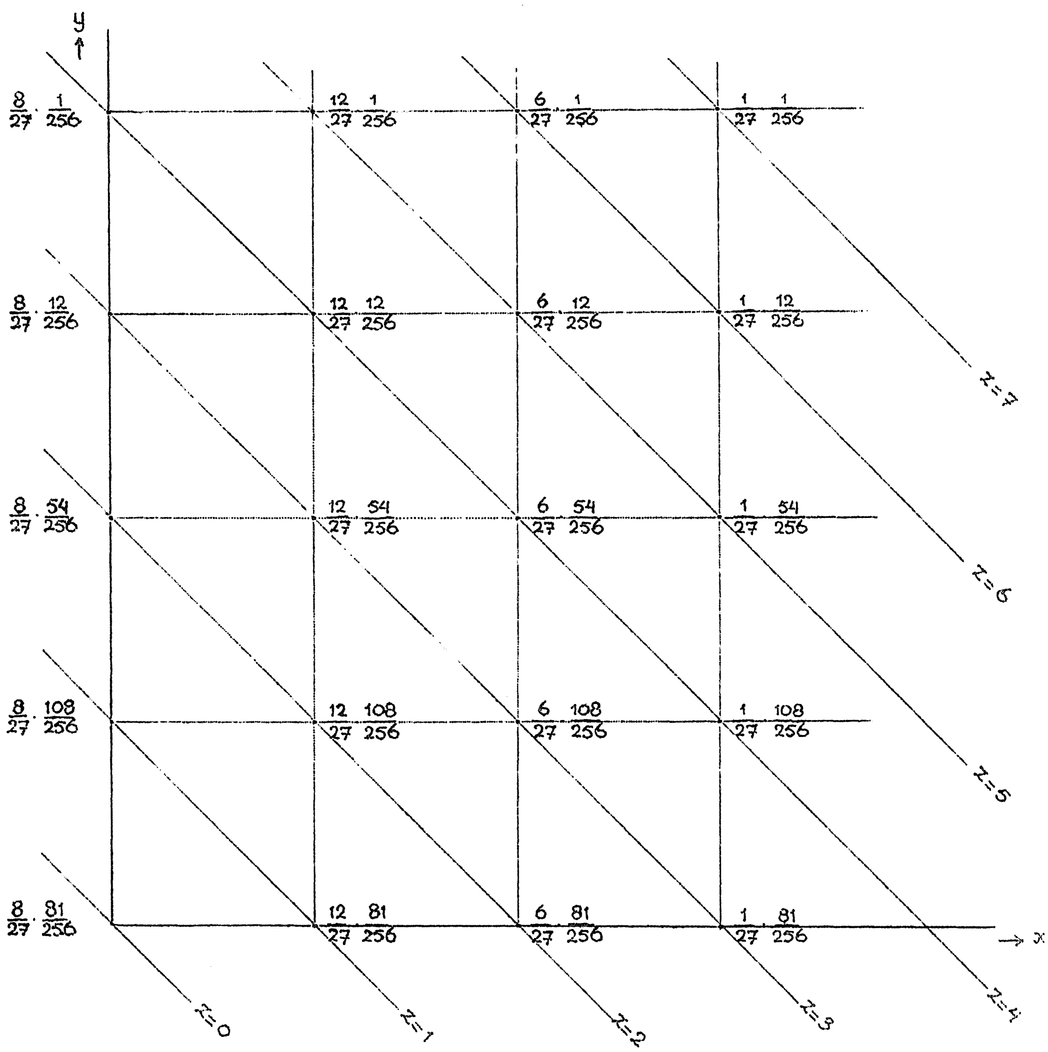
Voor $n=3$ en $p=\frac{1}{3}$ volgt uit (1):

$$P[\underline{x}=0]=\frac{8}{27}\,,\quad P[\underline{x}=1]=\frac{12}{27}\,,\quad P[\underline{x}=2]=\frac{6}{27}\,,\quad P[\underline{x}=3]=\frac{1}{27}\,.$$

Voor $m=4$ en $p_1=\frac{1}{4}$ volgt uit (2):

$$P[\underline{y}=0]=\frac{81}{256}\,,\quad P[\underline{y}=1]=\frac{108}{256}\,,\quad P[\underline{y}=2]=\frac{54}{256}\,,\quad P[\underline{y}=3]=\frac{12}{256}\,,\quad P[\underline{y}=4]=\frac{1}{256}\,.$$

$\underline{x}$ en $\underline{y}$ bezitten een simultane wh-verdeling en zijn onderling onafhankelijk verdeeld. Deze verdeling van $\underline{x}$ en $\underline{y}$ kunnen wij voorstellen als een wh-verdeling op het (x,y)vlak.



De getallen bij ieder punt (x, y) in dit (x,y)vlak stellen voor $P[\underline{x}=x \text{ en } \underline{y}=y]$, welke immers voor onafhankelijk en discreet verdeelde $\underline{x}$ en $\underline{y}$ gelijk zijn aan $P[\underline{x}=x].P[\underline{y}=y]$.

Zo is $P[\underline{x}=2 \text{ en } \underline{y}=3] = P[\underline{x}=2].P[\underline{y}=3] = \frac{6}{27}\cdot\frac{12}{256}$.

De wh-verdeling van $\underline{z}=\underline{x}+\underline{y}$ wordt nu gevonden uit:

$$P[\underline{z}=z] = \sum_{x_i+y_j=z} P[\underline{x}=x_i]\,P[\underline{y}=y_j].$$

In de figuur zijn dit de whn bij de punten op de rechten $x+y=z$.
Dus:

$$P[\underline{z}=0] = \frac{81}{256}\cdot\frac{8}{27} = \frac{648}{6912},$$

$$P[\underline{z}=1] = \frac{108}{256}\cdot\frac{8}{27} + \frac{81}{256}\cdot\frac{12}{27} = \frac{1836}{6912},$$

$$P[\underline{z}=2] = \frac{54}{256}\cdot\frac{8}{27} + \frac{108}{256}\cdot\frac{12}{27} + \frac{81}{256}\cdot\frac{6}{27} = \frac{2214}{6912},$$

$$P[\underline{z}=3] = \frac{12}{256}\cdot\frac{8}{27} + \frac{54}{256}\cdot\frac{12}{27} + \frac{108}{256}\cdot\frac{6}{27} + \frac{81}{256}\cdot\frac{1}{27} = \frac{1473}{6912},$$

$$P[\underline{z}=4] = \frac{1}{256}\cdot\frac{8}{27} + \frac{12}{256}\cdot\frac{12}{27} + \frac{54}{256}\cdot\frac{6}{27} + \frac{108}{256}\cdot\frac{1}{27} = \frac{584}{6912},$$

$$P[\underline{z}=5] = \frac{1}{256}\cdot\frac{12}{27} + \frac{12}{256}\cdot\frac{6}{27} + \frac{54}{256}\cdot\frac{1}{27} = \frac{138}{6912},$$

$$P[\underline{z}=6] = \frac{1}{256}\cdot\frac{6}{27} + \frac{12}{256}\cdot\frac{1}{27} = \frac{18}{6912},$$

$$P[\underline{z}=7] = \frac{1}{256}\cdot\frac{1}{27} = \frac{1}{6912}.$$

(Controle: de som van deze whn is gelijk aan 1).
Het gemiddelde en de spreiding van $\underline{z}$ kunnen het gemakkelijkst berekend worden uit:

$$\mathcal{E}\underline{z} = \mathcal{E}\underline{x} + \mathcal{E}\underline{y},$$

waarin

$$\mathcal{E}\underline{x} = 3.\frac{1}{3} = 1,$$

$$\mathcal{E}\underline{y} = 4.\frac{1}{4} = 1,$$

dus

$$\mathcal{E}\underline{z} = 2$$

en:

$$\sigma^2\{\underline{z}\} = \sigma^2\{\underline{x}\} + \sigma^2\{\underline{y}\},$$

want $\underline{x}$ en $\underline{y}$ zijn stochastisch onafhankelijk. Dus:

$$\sigma^2\{\underline{x}\} = 3.\frac{1}{3}\cdot\frac{2}{3} = \frac{2}{3},$$

$$\sigma^2\{\underline{y}\} = 4.\frac{1}{4}\cdot\frac{3}{4} = \frac{3}{4}$$

en

$$\sigma^2\{z\} = \frac{17}{12},$$

$$\sigma\{z\} = 1,19.$$

48. Met de toets van Wilcoxon toetsen wij de hypothese, H_0, dat de garens I en II even sterk zijn. Deze toets voeren wij uit zowel met de waarnemingen der trekproeven van de met A behandelde, als de niet met A behandelde garens. Wij krijgen het volgende resultaat:

I zonder A (x): 11,40 12,10 12,66 13,20 13,37;
II zonder A (y): 10,19 10,34 10,69 11,46 11,79;

$$yyyxyyxxxx;$$

$$u = 23.$$

Volgens tabel 5 van de éénzijdige kritieke waarden van u, zijn de kritieke waarden als $n = m = 5$ en tweezijdige $\alpha = 0,05$, 2 en 23.

Vervolgens voeren wij de toets van Wilcoxon uit met de gegevens van de met A behandelde garens.

I met A (x): 12,54 12,97 13,41 13,57 13,62;
II met A (y): 10,59 11,52 11,64 11,70 12,63;

$$yyyyxyxxxx;$$

$$u = 24.$$

Conclusie: Zowel voor de behandelde als de niet behandelde garens moeten wij de hypothese: de garens zijn even sterk, verwerpen ten gunste van de hypothese dat I sterker is dan II, behoudens een onbetrouwbaarheid 0,05.

Met de toets van Wilcoxon toetsen wij vervolgens de hypothese, H_0, dat het prepareren van de garens geen invloed heeft op de sterkte van de garens. Daar de sterkte van I en II verschilt, nemen wij de beide soorten garens niet tezamen, maar passen wij de toets op ieder van deze garens apart toe en combineren de gevonden resultaten door optelling der gevonden waarden van u.

I zonder A (x): 11,40 12,10 12,66 13,20 13,37;
I met A (y): 12,54 12,97 13,41 13,57 13,62;

$$xxyxyxxyyy;$$

$$u = 5.$$

II zonder A (x): 10,19 10,34 10,69 11,46 11,79;
II met A (y): 10,59 11,52 11,64 11,70 12,63;

$$xxyxxyyyxy\;;$$

$$U=6.$$

Voor garen I en II afzonderlijk kunnen wij de hypothese: toepassing van methode A heeft geen invloed op de sterkte van de garens, dus niet verwerpen. Vervolgens combineren wij de twee toetsen als volgt:

$$U_s \stackrel{def}{=} \sum_{i=1}^{k} U_i$$

met verwachting en variantie, onder H_0,

$$\mu_s \stackrel{def}{=} \frac{1}{2}\sum_{i=1}^{k} m_i n_i$$

en

$$\sigma_s^2 \stackrel{def}{=} \frac{1}{12}\sum_{i=1}^{k} m_i n_i (m_i+n_i+1).$$

Nu is:

$$U_s = 5+6 = 11\;;$$

$$\mu_s = 2.\tfrac{1}{2}.5.5 = 25\;;$$

$$\sigma_s^2 = 2.\tfrac{1}{12}.5.5.11 = 45{,}83$$

dus

$$\sigma_s = 6{,}77.$$

De overschrijdingskans, vinden wij door in tabel 1 van de $N(0,1)$-verdeling de overschrijdingskans op te zoeken van:

$$u = \frac{|11-25|-\frac{1}{2}}{6{,}77} = 1{,}99.$$

Deze bedraagt

$$k = 2 \times 0{,}0233 = 0{,}047.$$

Conclusie: Op grond van dit resultaat moeten wij de hypothese H_0 verwerpen ten gunste van de hypothese dat toepassing van methode A de garens sterker maakt. Het is dus aan te raden de methode A in te voeren.

Opmerking: Indien het invoeren van deze methode grote kosten met zich meebrengt, verdient het wellicht aanbeveling een grotere proef te verrichten, om met grotere stelligheid tot de genoemde conclusie te kunnen komen.

49. De statistische grootheid:

$$\underline{S}^2 = \frac{1}{k-1} \sum_{i=1}^{k} \{ \underline{x}_i - \underline{\bar{x}} \}^2$$

is een zuivere schatting van σ^2, als $\underline{x}_1, \ldots, \underline{x}_k$ onderling onafhankelijk verdeeld zijn volgens dezelfde wh-verdeling. Dit geldt algemeen, dus ook als die wh-verdeling binomiaal is. Derhalve is

$$\mathcal{E} \frac{1}{k-1} \sum_{i=1}^{k} \{ \underline{x}_i - \underline{\bar{x}} \}^2 = \sigma^2 = npq,$$

$$\mathcal{E} \frac{1}{k} \sum_{i=1}^{k} \{ \underline{x}_i - \underline{\bar{x}} \}^2 = \frac{k-1}{k} npq.$$

50. De hypothese H_0, die wij wensen te toetsen luidt: De kans op slagen is voor de studenten niet afhankelijk van hun richting.
Wij toetsen H_0 met de toets van Wilcoxon. De ene steekproef bestaat uit de resultaten der A-studenten en de andere uit die der B-studenten.

Het resultaat van het tentamen noteren wij in onderstaande z.g. 2 x 2 -tabel:

Kenmerk	A	B	totaal
0	11	16	27
1	17	52	69
totaal	28	68	96

Voor ieder der gezakte A-studenten geldt dat 16 B-studenten hetzelfde resultaat (= ook gezakt) hebben. De bijdrage tot $\underline{U}$ voor de 11 A-studenten is dus $\frac{1}{2}.11 \times 16$.
52 B-studenten hebben een beter resultaat dan elk der 11 gezakte A-studenten. De bijdrage tot $\underline{U}$ van deze groep is 11 x 52. Voor ieder der geslaagde A-studenten geldt dat 52 B-studenten hetzelfde resultaat hadden. De bijdrage tot $\underline{U}$ is dus voor deze 17 studenten $\frac{1}{2}.52 \times 17$. Voor $\underline{U}$ vinden wij dus:

$$U = \tfrac{1}{2}.11.16 + 11.52 + \tfrac{1}{2}.17.52 = 1102$$

$$\mu = \tfrac{1}{2} nm = \tfrac{1}{2}.28.68 = 952.$$

De spreiding van de verdeling van $\underline{U}$ wordt, nu er gelijken optreden, gevonden uit de volgende formule:

$$\sigma\{\underline{U}|t_1,\ldots,t_n, H_0\} = \sqrt{\tfrac{1}{12}nm(n+m+1) - \frac{nm}{12(m+n)(m+n-1)}\sum_{i=1}^{H}(t_i^3 - t_i)},$$

waarin de t_i de uitgebreidheden der groepen gelijken voorstellen. Hier zijn 2 groepen gelijken, één van 27 en één van 69, dus:

$$\sigma\{\underline{U}|t_1=27, t_2=69, H_0\} = \sqrt{\tfrac{1}{12}.28.68.97 - \frac{28.68}{12.96.95}\{(27^3-27)+(69^3-69)\}} =$$

$$= \sqrt{9334,61} = 96,6.$$

Onder H_0 is $\underline{U}$ bij benadering normaal verdeeld met gemiddelde $\mu = 952$ en spreiding $\sigma = 96,6$ en de overschrijdingskans kan dus gevonden worden (bij benadering) uit tabel 1 van de $N(0,1)$-verdeling.

De continuïteitscorrectie, die wij moeten toepassen (zie de opmerking bij opgave 17) is gelijk aan de helft van de afstand tussen twee opeenvolgende waarden van $\underline{U}$. Deze afstand is in dit geval constant en gelijk aan $\frac{1}{2}(n+m)$, zoals verderop wordt bewezen, zodat de continuïteitscorrectie dus $\frac{1}{4}(n+m)$ wordt en de tweezijdige overschrijdingskans bij benadering gelijk is aan twee maal het oppervlak van de $N(0,1)$-verdeling rechts van:

$$u = \frac{|U-\mu| - \frac{1}{4}(n+m)}{\sigma} = \frac{|1102-952|-24}{96,6} = 1,30$$

$$k = 0,19.$$

Dat de afstand tussen twee opeenvolgende waarden van U gelijk is aan $\frac{1}{2}(n+m)$ kunnen wij als volgt afleiden:

Het resultaat van het tentamen kan in het algemeen worden genoteerd als volgt:

Kenmerk	A	B	
0	a	c	t_1
1	b	d	t_2
	n	m	N

n, m, t_1 en t_2 zijn de randtotalen, $n+m = t_1+t_2 = N$.
De U van Wilcoxon is nu gelijk aan:

$$U = \tfrac{1}{2}ac + ad + \tfrac{1}{2}bd,$$

$$2U = ac + 2ad + bd,$$

met

$$\mu = \mathcal{E}\underline{U} = \tfrac{1}{2}nm.$$

Verder is

$$2(U-\mu) = ac + 2ad + bd - nm.$$

Dit kan als volgt herleid worden:

$$\begin{aligned} 2(U-\mu) &= a(a+b+c+d) - a(a+b) + ad + bd - nm = \\ &= aN - an + nd - nm = \\ &= aN - n(a-d+c+d) = \\ &= aN - nt_1. \end{aligned}$$

Hierin is alleen a variabel, daar de toets voorwaardelijk wordt uitgevoerd bij vaste waarden van N, n en t_1. Daar de afstand tussen 2 opeenvolgende waarden, die a aan kan nemen, gelijk is aan 1, liggen dus 2 opeenvolgende waarden van $2(U-\mu)$ op een afstand N van elkaar. Dit geldt dus ook voor $2U$, aangezien de ~~gereduceerde~~ gereduceerde grootheid $U-\mu$ en de grootheid U een bedrag μ t.o.v. elkaar zijn verschoven. De afstand tussen 2 opeenvolgende waarden van U is dus $\frac{1}{2}N$. De continuïteitscorrectie bedraagt derhalve $\frac{1}{4}N$.

Passen wij geen continuïteitscorrectie toe dan vinden wij voor de tweezijdige overschrijdingskans twee maal het oppervlak van de normale verdeling rechts van

$$u = \frac{|U-\mu|}{\sigma} = \frac{|1102-952|}{96{,}6} = 1{,}55$$

$$k = 0{,}12.$$

Deze overschrijdingskans is dus wel aanmerkelijk kleiner dan die welke gevonden is bij toepassing van de continuïteitscorrectie en het is dus niet raadzaam deze weg te laten.

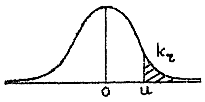
De conclusie luidt:

De hypothese H_0, dat de kans op slagen gelijk is voor beide richtingen, kan niet verworpen worden.

Tabel 1.

Tabel van de N(0,1)-verdeling.

Deze tabel bevat de rechteroverschrijdingskansen, vermenigvuldigd met 10^4, van een N(0,1)-verdeelde stochastische grootheid $\underline{u}$ d.w.z.



$k_\tau = P[\underline{u} > u] = P[\underline{u} \geq u] = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-\frac{1}{2}x^2} dx.$

u	0	1	2	3	4	5	6	7	8	9
0,0	5000	4960	4920	4880	4840	4801	4761	4721	4681	4641
0,1	4602	4562	4522	4483	4443	4404	4364	4325	4286	4247
0,2	4207	4168	4129	4090	4052	4013	3974	3936	3897	3859
0,3	3821	3783	3745	3707	3669	3632	3594	3557	3520	3483
0,4	3446	3409	3372	3336	3300	3264	3228	3192	3156	3121
0,5	3085	3050	3015	2981	2946	2912	2877	2843	2810	2776
0,6	2743	2709	2676	2643	2611	2578	2546	2514	2483	2451
0,7	2420	2389	2358	2327	2296	2266	2236	2206	2177	2148
0,8	2119	2090	2061	2033	2005	1977	1949	1922	1894	1867
0,9	1841	1814	1788	1762	1736	1711	1685	1660	1635	1611
1,0	1587	1562	1539	1515	1492	1469	1446	1423	1401	1379
1,1	1357	1335	1314	1292	1271	1251	1230	1210	1190	1170
1,2	1151	1131	1112	1093	1075	1056	1038	1020	1003	0985
1,3	0968	0951	0934	0918	0901	0885	0869	0853	0838	0823
1,4	0808	0793	0778	0764	0749	0735	0721	0708	0694	0681
1,5	0668	0655	0643	0630	0618	0606	0594	0582	0571	0559
1,6	0548	0537	0526	0516	0505	0495	0485	0475	0465	0455
1,7	0446	0436	0427	0418	0409	0401	0392	0384	0375	0367
1,8	0359	0351	0344	0336	0329	0322	0314	0307	0301	0294
1,9	0287	0281	0274	0268	0262	0256	0250	0244	0239	0233
2,0	0228	0222	0217	0212	0207	0202	0197	0192	0188	0183
2,1	0179	0174	0170	0166	0162	0158	0154	0150	0146	0143
2,2	0139	0136	0132	0129	0125	0122	0119	0116	0113	0110
2,3	0107	0104	0102	0099	0096	0094	0091	0089	0087	0084
2,4	0082	0080	0078	0075	0073	0071	0069	0068	0066	0064
2,5	0062	0060	0059	0057	0055	0054	0052	0051	0049	0048
2,6	0047	0045	0044	0043	0041	0040	0039	0038	0037	0036
2,7	0035	0034	0033	0032	0031	0030	0029	0028	0027	0026
2,8	0026	0025	0024	0023	0023	0022	0021	0021	0020	0019
2,9	0019	0018	0018	0017	0016	0016	0015	0015	0014	0014
3,0	0013	0013	0013	0012	0012	0011	0011	0011	0010	0010
3,1	0010	0009	0009	0009	0008	0008	0008	0008	0007	0007
3,2	0007	0007	0006	0006	0006	0006	0006	0005	0005	0005
3,3	0005	0005	0005	0004	0004	0004	0004	0004	0004	0003
3,4	0003	0003	0003	0003	0003	0003	0003	0003	0003	0002

Voorbeelden: 1) $P[\underline{u} \geq 1{,}73] = 0{,}0418$; $P[\underline{u} < 0{,}52] = 1 - P[\underline{u} \geq 0{,}52] = 1 - 0{,}3015 = 0{,}6985$.

2) De waarde van u waarvoor $P[\underline{u} \geq u] = 0{,}05$ vindt men door "terugzoeken" in de tabel. Deze blijkt $u = 1{,}645$ te zijn (met lineaire interpolatie).

Tabel 2.

Betrouwbaarheidsgrenzen voor $n\theta$ voor $x \leqq 10$ en grote n.
(α éénzijdig)

x \ α	0,005		0,025		0,05		0,10	
	$n\theta_*$	$n\theta^*$	$n\theta_*$	$n\theta^*$	$n\theta_*$	$n\theta^*$	$n\theta_*$	$n\theta^*$
0	0	5,3	0	3,7	0	3,0	0	2,3
1	0,005	7,4	0,025	5,6	0,05	4,7	0,11	3,9
2	0,1	9,3	0,2	7,2	0,4	6,3	0,5	5,3
3	0,3	11,0	0,6	8,8	0,8	7,8	1,1	6,7
4	0,6	12,6	1,0	10,2	1,4	9,2	1,7	8,0
5	1,0	14,1	1,6	11,7	2,0	10,5	2,4	9,3
6	1,5	15,6	2,2	13,1	2,6	11,8	3,1	10,5
7	2,0	17,1	2,8	14,4	3,3	13,1	3,9	11,8
8	2,5	18,5	3,4	15,8	4,0	14,4	4,7	13,0
9	3,1	20,0	4,0	17,1	4,7	15,7	5,4	14,2
10	3,7	21,3	4,7	18,4	5,4	17,0	6,2	15,4

Deze tabel bevat betrouwbaarheidsgrenzen voor $n\theta$, waarin θ een onbekende fractie of wh voorstelt en n het aantal onafhankelijke waarnemingen, dat verricht is om θ te schatten (de omvang van de steekproef). Is x het aantal successen, dat bij deze n waarnemingen verkregen is en is α de (éénzijdige) onbetrouwbaarheidsdrempel, die men gekozen heeft, dan leest men in de bijbehorende regel en kolom de onder- en bovengrens voor $n\theta$ af. Met een onbetrouwbaarheidsdrempel α geldt dan

$$n\theta_* \leqq n\theta \quad \text{dus} \quad \theta_* \leqq \theta$$

en eveneens met onbetrouwbaarheidsdrempel α

$$n\theta^* \geqq n\theta \quad \text{dus} \quad \theta^* \geqq \theta,$$

terwijl de uitspraak

$$\theta_* \leqq \theta \leqq \theta^*$$

een onbetrouwbaarheidsdrempel 2α heeft. De gelijkheidstekens mogen ook weggelaten worden.

Voorbeelden:

1) Is $n=150$, $x=4$ en α(éénzijdig)$=0{,}025$ dan is $\theta_* = \frac{1}{150} = 0{,}0066$ en $\theta^* = \frac{10{,}2}{150} = 0{,}068$.

2) Als bij $n=100$ en $\alpha=0{,}05$ gevraagd wordt de grootste x te vinden, waarbij nog $\theta^* \leqq 0{,}12$ is, dan moet dus $n\theta^* = 100\theta^* \leqq 12$ zijn. Dit is voor $x=6$ ($n\theta^* = 11{,}8$) nog het geval, voor $x=7$ ($n\theta^* = 13{,}1$) niet meer. De oplossing is dus $x=6$. De tabel berust op de Poisson-benadering van de binomiale verdeling. De benadering is goed als $x \ll n$. Is $n-x \ll n$ en $n-x \leqq 10$ dan kan men $n-x$ in plaats van x gebruiken. Men vindt dan op dezelfde wijze betrouwbaarheidsgrenzen voor $1-\theta$.

Toelichting bij de tabel van de tekentoets.

Deze tabel geeft de grenzen van de linker-kritieke zones van de tweezijdige tekentoets, waarbij n het aantal stochastisch onafhankelijke waarnemingen voorstelt. De onbetrouwbaarheidsdrempels zijn tweezijdig opgegeven. Voor éénzijdige toetsing moeten zij dus gehalveerd worden. De grens van de rechterkritieke zone wordt uit de in de tabel opgegeven waarde verkregen door deze van n af te trekken.

Voorbeelden:

1) $n=87$, $x=32$ (x = aantal successen): deze uitkomst valt in de linkerhelft van de kritieke zone met onbetrouwbaarheidsdrempel 0,02. Is dus α (tweezijdig) $\geq 0{,}02$ gekozen, dan wordt de hypothese H_0 dat de kans op een succes $= \frac{1}{2}$ is, verworpen ten gunste van de hypothese, dat deze kans $< \frac{1}{2}$ is. Bij links-éénzijdige toetsing is dit ook nog zo met $\alpha = 0{,}01$; bij rechtséénzijdige toetsing wordt H_0 niet verworpen. Is α (tweezijdig) $= 0{,}01$, dan wordt H_0 niet verworpen.

2) $n=54$, $x=34$, α (tweezijdig) $= 0{,}05$. De grens van de rechterhelft van de kritieke zone is gelijk aan $54 - 19 = 35$. Hier valt x niet in, dus kan H_0 niet verworpen worden. Uit de tabel blijkt tevens, dat voor α (tweezijdig) $= 0{,}10$, x juist in de kritieke zone ligt. Bij tweezijdige toetsing met $\alpha = 0{,}10$ of rechts-éénzijdige met α(éénzijdig)$=0{,}05$ wordt H_0 dus nog juist verworpen, ten gunste van de hypothese, dat de kans op succes $> \frac{1}{2}$ is.

Tabel 3.

Kritieke waarden van de tweezijdige tekentoets.

n	Onbetrouwbaarheidsdrempel 0,01	0,02	0,05	0,10	n	Onbetrouwbaarheidsdrempel 0,01	0,02	0,05	0,10
1	-	-	-	-	51	15	16	18	19
2	-	-	-	-	52	16	17	18	19
3	-	-	-	-	53	16	17	18	20
4	-	-	-	-	54	17	18	19	20
5	-	-	-	0	55	17	18	19	20
6	-	-	0	0	56	17	18	20	21
7	-	0	0	0	57	18	19	20	21
8	0	0	0	1	58	18	19	21	22
9	0	0	1	1	59	19	20	21	22
10	0	0	1	1	60	19	20	21	23
11	0	1	1	2	61	20	20	22	23
12	1	1	2	2	62	20	21	22	24
13	1	1	2	3	63	20	21	23	24
14	1	2	2	3	64	21	22	23	24
15	2	2	3	3	65	21	22	24	25
16	2	2	3	4	66	22	23	24	25
17	2	3	4	4	67	22	23	25	26
18	3	3	4	5	68	22	23	25	26
19	3	4	4	5	69	23	24	25	27
20	3	4	5	5	70	23	24	26	27
21	4	4	5	6	71	24	25	26	28
22	4	5	5	6	72	24	25	27	28
23	4	5	6	7	73	25	26	27	28
24	5	5	6	7	74	25	26	28	29
25	5	6	7	7	75	25	26	28	29
26	6	6	7	8	76	26	27	28	30
27	6	7	7	8	77	26	27	29	30
28	6	7	8	9	78	27	28	29	31
29	7	7	8	9	79	27	28	30	31
30	7	8	9	10	80	28	29	30	32
31	7	8	9	10	81	28	29	31	32
32	8	8	9	10	82	28	30	31	33
33	8	9	10	11	83	29	30	32	33
34	9	9	10	11	84	29	30	32	33
35	9	10	11	12	85	30	31	32	34
36	9	10	11	12	86	30	31	33	34
37	10	10	12	13	87	31	32	33	35
38	10	11	12	13	88	31	32	34	35
39	11	11	12	13	89	31	33	34	36
40	11	12	13	14	90	32	33	35	36
41	11	12	13	14	91	32	33	35	37
42	12	13	14	15	92	33	34	36	37
43	12	13	14	15	93	33	34	36	38
44	13	13	15	16	94	34	35	37	38
45	13	14	15	16	95	34	35	37	38
46	13	14	15	16	96	34	36	37	39
47	14	15	16	17	97	35	36	38	39
48	14	15	16	17	98	35	37	38	40
49	15	15	17	18	99	36	37	39	40
50	15	16	17	18	100	36	37	39	41

Toelichting bij de tabel van de toets van Student.

Is $x_1, \ldots, x_n$ een steekproef uit een $N(\mu,\sigma)$-verdeling met onbekende μ en σ en wenst men de hypothese $H_0 : \mu = \mu_0$ te toetsen, dan berekent men

$$t_1 = \frac{\bar{x}-\mu_0}{s}\sqrt{n} \quad (\bar{x} = \tfrac{1}{n}\Sigma x_i \; ; \; s^2 = \tfrac{1}{n-1}\Sigma(x_i-\bar{x})^2)$$

en zoekt in de tabel, in de regel met $\nu = n-1$ (ν = het aantal "vrijheidsgraden") de waarden op, waar t_1 tussen ligt. Aan de kop van de desbetreffende kolommen vindt men dan twee grenzen, waar de tweezijdige overschrijdingskans tussen ligt.

Voorbeeld: 1) $t_1 = 1{,}4$, $\nu = 15$. De tabel geeft $0{,}2 > k > 0{,}1$ waarin k tweezijdig is. H_0 wordt niet verworpen, als $\alpha \leqq 0{,}1$.

2) $t_1 = -2{,}3$, $\nu = 7$. De verdeling van Student is symmetrisch t.o.v. 0. Wij zoeken in de tabel, in de regel met $\nu = 7$, de getallen, waar + 2,3 tussen ligt. Dit geeft voor de tweezijdige overschrijdingskans de waarden 0,1 en 0,05. Met α (tweezijdig) = 0,05 wordt H_0 dus niet verworpen. Bij linkséénzijdige toetsing met $\alpha = 0{,}05$ zou wel tot verwerping van H_0 overgegaan worden en wel ten gunste van de hypothese $\mu < \mu_0$.

3) $\nu = 23$. Gevraagd de grenzen van de tweezijdige kritieke zone van t_1 voor $\alpha = 0{,}01$. De kolom met $\alpha = 0{,}01$ geeft in de regel met $\nu = 23$ de waarde 2,807. De grenzen zijn dus - 2,807 en + 2,807.

Zijn $x_1, \ldots, x_n$ en $y_1, \ldots, y_m$ twee onafhankelijke steekproeven uit 2 normale verdelingen met μ_1 resp. μ_2 als gemiddelden en gelijke σ, en wenst men de hypothese

$$H_0 : \mu_1 - \mu_2 = \Delta$$

te toetsen, dan berekent men

$$t_2 = \frac{\bar{x}-\bar{y}-\Delta}{\sqrt{S_1^2+S_2^2}}\sqrt{\frac{mn(m+n-2)}{m+n}},$$

waarin $S_1^2 = \Sigma(x_i-\bar{x})^2$ en $S_2^2 = \Sigma(y_j-\bar{y})^2$ en gaat verder te werk als boven, met $\nu = n+m-2$.

Tabel 4.

Tweezijdige overschrijdingskansen van de toetsen van Student.

ν	k = .9	.8	.7	.6	.5	.4	.3	.2	.1	.05	.02	.01
1	.158	.325	.510	.727	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657
2	.142	.289	.445	.617	.816	1.061	1,386	1.886	2.920	4.303	6.965	9.925
3	.137	.277	424	584	.765	.978	1.250	1.638	2.353	3.182	4.541	5.841
4	.134	.271	.414	.569	.741	.941	1.190	1.533	2.132	2.776	3.747	4.604
5	.132	.267	.408	.559	.727	.920	1.156	1.476	2.015	2.571	3.365	4.032
6	.131	.265	.404	.553	.718	.906	1.134	1.440	1.943	2.447	3.143	3.707
7	.130	.263	.402	.549	.711	.896	1.119	1.415	1.895	2.365	2.998	3.499
8	.130	.262	.399	.546	.706	.889	1.108	1.397	1.860	2.306	2.896	3.355
9	.129	.261	.398	.543	.703	.883	1.100	1.383	1.833	2.262	2.821	3.250
10	.129	.260	.397	.542	.700	.879	1.093	1.372	1.812	2.228	2.764	3.169
11	.129	.260	.396	.540	.697	.876	1.088	1.363	1.796	2.201	2.718	3.106
12	.128	.259	.395	.539	.695	.873	1.083	1.356	1.782	2.179	2.681	3.055
13	.128	.259	.394	.538	.694	.870	1.079	1.350	1.771	2.160	2.650	3.012
14	.128	.258	.393	.537	.692	.868	1.076	1.345	1.761	2.145	2.624	2.977
15	.128	.258	.393	.536	.691	.866	1.074	1.341	1.753	2.131	2.602	2.947
16	.128	.258	.392	.535	.690	.865	1.071	1.337	1.746	2.120	2.583	2.921
17	.128	.257	.392	.534	.689	.863	1.069	1.333	1.740	2.110	2.567	2.898
18	.127	.257	.392	.534	.688	.862	1.067	1.330	1.734	2.101	2.552	2.878
19	.127	.257	.391	.533	.688	.861	1.066	1.328	1.729	2.093	2.539	2.861
20	.127	.257	.391	.533	.687	.860	1.064	1.325	1.725	2.086	2.528	2.845
21	.127	.257	.391	.532	.686	.859	1.063	1.323	1.721	2.080	2.518	2.831
22	.127	.256	.390	.532	.686	.858	1.061	1.321	1.717	2.074	2.508	2.819
23	.127	.256	.390	.532	.685	.858	1.060	1.319	1.714	2.069	2.500	2.807
24	.127	.256	.390	.531	.685	.857	1.059	1.318	1.711	2.064	2.492	2.797
25	.127	.256	.390	.531	.684	.856	1.058	1.316	1.708	2.060	2.485	2.787
26	.127	.256	.390	.531	.684	.856	1.058	1.315	1.706	2.056	2.479	2.779
27	.127	.256	.389	.531	.684	.855	1.057	1.314	1.703	2.052	2.473	2.771
28	.127	.256	.389	.530	.683	.855	1.056	1.313	1.701	2.048	2.467	2.763
29	.127	.256	.389	.530	.683	.854	1.055	1.311	1.699	2.045	2.462	2.756
30	.127	.256	.389	.530	.683	.854	1.055	1.310	1.697	2.042	2.457	2.750
∞	.12566	.25335	.38532	.52440	.67449	.84162	1.03643	1.28155	1.64485	1.95996	2.32634	2.57582

Toelichting bij de tabel voor de toets van Wilcoxon.

Deze tabel bevat de grenzen van de linker- en van de rechterhelft van de kritieke zone van de toets van Wilcoxon voor 3 waarden van α . De omvang van de beide steekproeven is m resp. n . Deze twee mogen in de tabel verwisseld worden. In de linker-onderhelft van ieder tabelletje vindt men de grenzen der linker-kritieke zones en in de rechterbovenhelften die van de rechter-kritieke zones.

Voorbeeld. $m=4$, $n=8$, α (tweezijdig) $= 0{,}05$. In de regel voor $m=4$ en de kolom voor $n=8$ van de tabel met α (éénzijdig) $= 0{,}025$ vinden wij 28. Daar $mn = 32$ is dit de grens van de rechterhelft van de tweezijdige kritieke zone. De grens van de linkerhelft is $32 - 28 = 4$. Dit getal vindt men ook in dezelfde tabel in de regel voor $m=8$ en de kolom voor $n=4$.

Tabel 5.

Eénzijdig kritieke waarden van de toets van Wilcoxon($m,n \leq 10$).

U_α \ $mn - U_\alpha$

α = 0,05(éénzijdig)

m \ n	3	4	5	6	7	8	9	10
3	0 \ 9	12	14	16	19	21	23	26
4	0	1 \ 15	18	21	24	27	30	33
5	1	2	4 \ 21	25	29	32	36	39
6	2	3	5	7 \ 29	34	38	42	46
7	2	4	6	8	11 \ 38	43	48	53
8	3	5	8	10	13	15 \ 49	54	60
9	4	6	9	12	15	18	21 \ 60	66
10	4	7	11	14	17	20	24	27 \ 73

α = 0,025(éénzijdig)

m \ n	3	4	5	6	7	8	9	10
3	- \ -	-	15	17	20	22	25	27
4	-	0 \ 16	19	22	25	28	32	35
5	0	1	2 \ 23	27	30	34	38	42
6	1	2	3	5 \ 31	36	40	44	49
7	1	3	5	6	8 \ 41	46	51	56
8	2	4	6	8	10	13 \ 51	57	63
9	2	4	7	10	12	15	17 \ 64	70
10	3	5	8	11	14	17	20	23 \ 77

α= 0,01(éénzijdig)

m \ n	3	4	5	6	7	8	9	10
3	- \ -	-	-	-	21	24	26	29
4	-	- \ -	20	23	27	30	33	37
5	-	0	1 \ 24	28	32	36	40	44
6	-	1	2	3 \ 33	38	42	47	52
7	0	1	3	4	6 \ 43	49	54	59
8	0	2	4	6	7	9 \ 55	61	67
9	1	3	5	7	9	11	14 \ 67	74
10	1	3	6	8	11	13	16	19 \ 81

Tabel 6.

Aselecte getallen 0,...,9

rij \ kolom	1 t/m 10	11 t/m 20	21 t/m 30	31 t/m 40
1	09 18 82 00 97	32 82 53 95 27	04 22 08 63 04	83 33 98 73 74
	90 04 58 54 97	51 98 15 06 54	94 93 88 19 97	91 87 07 61 50
t/m	73 18 95 02 07	47 67 72 62 69	62 29 06 44 64	27 12 46 70 18
	75 76 87 64 90	20 97 18 17 49	90 42 91 22 72	95 37 50 58 71
5	54 01 64 40 56	66 28 13 10 03	00 68 22 73 98	20 71 45 32 95
6	08 35 86 99 10	78 54 24 27 85	13 66 15 88 73	04 61 89 75 53
	28 30 60 32 64	81 33 31 05 91	40 51 00 78 93	32 60 46 04 75
t/m	53 84 08 62 33	81 59 41 36 28	51 21 59 02 90	28 46 66 87 95
	91 75 75 37 41	61 61 36 22 69	50 26 39 02 12	55 78 17 65 14
10	89 41 59 26 94	00 39 75 83 91	12 60 71 76 46	48 94 97 23 06
11	77 51 30 38 20	86 83 42 99 01	68 41 48 27 74	51 90 81 39 80
	19 50 23 71 74	69 97 92 02 88	55 21 02 97 73	74 28 77 52 51
t/m	21 81 85 93 13	93 27 88 17 57	05 68 67 31 56	07 08 28 50 46
	51 47 46 64 99	68 10 72 36 21	94 04 99 13 45	42 83 60 91 91
15	99 55 96 83 31	62 53 52 41 70	69 77 71 28 30	74 81 97 81 42
16	33 71 34 80 07	93 58 47 28 69	51 92 66 47 21	58 30 32 98 22
	85 27 48 68 93	11 30 32 92 70	28 83 43 41 37	73 51 59 04 00
t/m	84 13 38 96 40	44 03 55 21 66	73 85 27 00 91	61 22 26 05 61
	56 73 21 62 34	17 38 59 61 31	10 12 19 16 22	85 49 65 75 60
20	65 13 85 68 06	87 64 88 52 61	34 31 36 58 61	45 87 52 10 69
21	38 00 10 21 76	81 71 91 17 11	71 60 29 29 37	74 21 96 40 49
	37 40 29 63 97	01 30 47 75 86	56 27 11 00 86	47 32 46 26 05
t/m	97 12 54 03 48	37 08 33 14 17	21 81 53 92 50	75 23 76 20 47
	21 82 64 11 34	47 14 33 40 72	64 63 88 59 02	49 13 90 64 41
25	73 13 54 27 42	95 71 [illegible] 90 35	85 79 47 42 96	08 78 98 81 56
26	07 63 87 79 29	03 06 11 80 72	96 20 74 41 56	23 82 19 95 38
	60 52 88 34 41	07 95 41 98 14	59 17 52 06 95	05 53 35 21 39
t/m	83 59 63 56 55	06 95 89 29 83	05 12 80 97 19	77 43 35 37 83
	10 85 06 27 46	99 59 91 05 07	13 49 90 63 19	53 07 57 18 39
30	39 82 09 89 52	43 62 26 31 47	64 42 18 08 14	43 80 00 93 51
31	59 58 00 64 78	75 56 97 88 00	88 83 55 44 86	23 76 80 61 56
	38 50 80 73 41	23 79 34 87 63	90 82 29 70 22	17 71 90 42 07
t/m	30 69 27 06 68	94 68 81 61 27	56 19 68 00 91	82 06 76 34 00
	65 44 39 56 59	18 28 82 74 37	49 63 22 40 41	08 33 76 56 76
35	27 26 75 02 64	13 19 27 22 94	07 47 74 46 06	17 98 54 89 11
36	91 30 70 69 91	19 07 22 42 10	36 69 95 37 28	28 82 53 57 93
	68 43 49 46 88	84 47 31 36 22	62 12 69 84 08	12 84 38 25 90
t/m	48 90 81 58 77	54 74 52 45 91	35 70 00 47 54	83 82 45 26 92
	06 91 34 51 97	42 67 27 86 01	11 88 30 95 28	63 01 19 89 01
40	10 45 51 60 19	14 21 03 37 12	91 34 23 78 21	88 32 58 08 51
41	12 88 39 73 43	65 02 76 11 84	04 28 50 13 92	17 97 41 50 77
	21 77 83 09 76	38 80 73 69 61	31 64 94 20 96	63 28 10 20 23
t/m	19 52 35 95 15	65 12 25 96 59	86 28 36 82 58	69 57 21 37 98
	67 24 55 26 70	35 58 31 65 63	79 24 68 66 86	76 46 33 42 22
45	60 58 44 73 77	07 50 03 79 92	45 13 42 65 29	26 76 08 36 37

Deze tabel bevat 1800 onafhankelijke aselecte trekkingen met teruglegging uit een populatie, waarin de getallen 0,1,...,9 even vaak voorkomen. Voor de overzichtelijkheid zijn zij in groepjes van twee gerangschikt. De kolommen en de rijen zijn genummerd om verwijzing gemakkelijker te maken.