STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

Rapport S 284

Contrôlegrenzen voor $\frac{\underline{w}}{\underline{x}}$ bij duplo-waarnemingen

door

W.R. van Zwet

1961

Contrôlegrenzen voor $\frac{\underline{w}}{\underline{x}}$ bij duplo-waarnemingen.

In het onderstaande wordt het oorspronkelijk gestelde probleem voor een analyse in duplo behandeld:
Bij een chemische analyse in duplo wordt verondersteld, dat de beide analyse-resultaten $\underline{x}_1$ en $\underline{x}_2$ onafhankelijk normaal verdeeld zijn met verwachting μ en variantie σ^2. Het wordt niet toelaatbaar geacht, dat de variatie-coëfficient $\frac{\sigma}{\mu}$ een bepaalde waarde overschrijdt. Gevraagd wordt hiertoe contrôlegrenzen te bepalen voor

$$\frac{\underline{w}}{\underline{x}} = 2.\frac{|\underline{x}_1 - \underline{x}_2|}{\underline{x}_1 + \underline{x}_2}.$$

Daar de analyse-resultaten begrensd zijn, $0 \leqslant \underline{x}_i \leqslant 1$ *), zal de veronderstelde normale verdeling van $\underline{x}_i$ slechts voor kleine waarden van $\frac{\sigma}{\mu}$ en $\frac{\sigma}{1-\mu}$ een bruikbare benadering van de werkelijke verdeling kunnen zijn. In het volgende zal dan ook worden aangenomen, dat

(1) $$\frac{\sigma}{\mu} \leqslant \frac{1}{3} \quad \text{en} \quad \frac{\sigma}{1-\mu} \leqslant \frac{1}{3},$$

waarin de waarde $\frac{1}{3}$ een vrij algemeen gebruikte, doch overigens willekeurige, grenswaarde is.
De gevraagde contrôlegrens g_α, die wordt gedefinieerd door

(2) $$P\left(\frac{\underline{w}}{\underline{x}} \geqslant g_\alpha\right) = \alpha.$$

kan niet op eenvoudige wijze exact worden berekend. Voor de kleine waarden van $\frac{\sigma}{\mu}$, die hier worden beschouwd, kan echter de volgende, zeer nauwkeurige benadering worden gevonden:

(3) $$g_\alpha = \frac{2\,u_{\alpha/2}}{\sqrt{2\left(\frac{\mu}{\sigma}\right)^2 - u^2_{\alpha/2}}}.$$

Hierin stelt $u_{\alpha/2}$ de waarde voor, die door een normaal verdeelde stochastische variabele met verwachting 0 en variantie 1 met waarschijnlijkheid $\alpha/2$ wordt overschreden:

(4) $$P(\underline{u} \geqslant u_{\alpha/2}) = \frac{1}{\sqrt{2\pi}} \int_{u_{\alpha/2}}^{\infty} e^{-\frac{1}{2}x^2}\,dx = \alpha/2. \qquad (\underline{u} \in N(0,1)).$$

Voor $\sigma/\mu \leqslant \frac{1}{3}$ is de fout in α tengevolge van deze benadering kleiner dan $2{,}2 \cdot 10^{-5}$ en dus voor de gebruikelijke waarden van α te verwaarlozen.
Voor een aantal waarden van $\frac{\sigma}{\mu} \leqslant 0{,}33$ en $\alpha = 0{,}10$; $0{,}05$; $0{,}025$; $0{,}01$ is g_α met behulp van (3) en (4) berekend en hieronder getabelleerd.
De fout in de opgegeven waarden van g_α ten gevolge van de gebruikte benadering is altijd kleiner dan 1 in de laatste decimaal.

*)Gedacht wordt aan een gehalte, uitgedrukt als verhouding.

Tabel: g_α als functie van $\frac{\sigma}{\mu}$ en α

$$P\left(\frac{\underline{w}}{\underline{\bar{x}}} \geqslant g_\alpha\right) = P\left(2 . \frac{|\underline{x}_1 - \underline{x}_2|}{\underline{x}_1 + \underline{x}_2} \geqslant g_\alpha\right) = \alpha$$

σ/μ \ α	0,10	0,05	0,025	0,01
0,005	0,012	0,014	0,016	0,018
0,01	0,023	0,028	0,032	0,036
0,02	0,047	0,055	0,063	0,073
0,03	0,070	0,083	0,095	0,109
0,04	0,093	0,111	0,127	0,146
0,05	0,117	0,139	0,159	0,183
0,06	0,140	0,167	0,191	0,220
0,08	0,187	0,223	0,256	0,295
0,10	0,234	0,280	0,321	0,370
0,15	0,354	0,425	0,490	0,568
0,20	0,478	0,577	0,668	0,782
0,25	0,608	0,739	0,863	1,023
0,30	0,745	0,914	1,081	1,305
0,33	0,831	1,029	1,227	1,504

Appendix   Afleiding van de resultaten

Onder de gemaakte veronderstellingen zijn $(\underline{x}_1 + \underline{x}_2)$ en $(\underline{x}_1 - \underline{x}_2)$ onafhankelijk en bezitten normale verdelingen met verwachtingen 2μ resp. 0 en variantie $2\sigma^2$. Dus geldt

$$\underline{t} = \frac{\underline{x}_1 + \underline{x}_2}{|\underline{x}_1 - \underline{x}_2|} = \frac{\frac{1}{\sigma\sqrt{2}}.(\underline{x}_1 + \underline{x}_2)}{\frac{1}{\sigma\sqrt{2}}.|\underline{x}_1 - \underline{x}_2|} = \frac{\underline{y}}{|\underline{v}|} = \frac{\underline{y}}{\underline{z}},$$

waarbij $\underline{y}$ en $\underline{v}$ onafhankelijk normale verdelingen bezitten met verwachtingen $\frac{\mu}{\sigma}\sqrt{2}$ resp. 0 en variantie 1 en $\underline{z} = |\underline{v}| \geqslant 0$.
Daar $\underline{z}^2 = \underline{v}^2$ dus een χ^2-verdeling met 1 vrijheidsgraad heeft, bezit

$$\underline{t}' = \frac{\underline{x}_1 + \underline{x}_2 - \frac{\mu}{\sigma}\sqrt{2}}{|x_1 - x_2|}$$

een Studentverdeling met 1 vrijheidsgraad en

$$\underline{t} = \frac{\underline{x}_1 + \underline{x}_2}{|\underline{x}_1 - \underline{x}_2|}$$

een niet-centrale t- verdeling met 1 vrijheidsgraad en niet-centraliteitsparameter $\delta = \frac{\mu}{\sigma}\sqrt{2}$. De gevraagde contrôlegrens $g_\alpha \geqslant 0$ met overschrijdingskans α wordt dus gegeven door

$$\alpha = P\left(\frac{\underline{w}}{\underline{\bar{x}}} \geqslant g_\alpha\right) = P\left(2.\frac{|\underline{x}_1 - \underline{x}_2|}{\underline{x}_1 + \underline{x}_2} \geqslant g_\alpha\right) =$$

$$= P\left(0 \leqslant \frac{\underline{x}_1 + \underline{x}_2}{|\underline{x}_1 - \underline{x}_2|} \leqslant \frac{2}{g_\alpha}\right) = P\left(0 \leqslant \underline{t}_{1;\,\frac{\mu}{\sigma}\sqrt{2}} \leqslant \frac{2}{g_\alpha}\right) \qquad (5)$$

waarin $\underline{t}_{1;\delta}$ een niet- centraal verdeelde Student-grootheid met 1 vrijheidsgraad en niet-centraliteitsparameter δ voorstelt.
Voor 1 vrijheidsgraad is de niet- centrale t - verdeling echter onvoldoende getabelleerd, zodat g_α niet zonder meer uit (5) kan worden bepaald. Voor kleine waarden van $\frac{\sigma}{\mu}$ kan evenwel een geschikte benadering worden gevonden.
De verdelingsdichtheid f(t) van de genoemde niet-centrale t-verdeling kan op eenvoudige wijze uit de verdelingsdichtheden

$$f_1(y) = \frac{1}{\sqrt{2\pi}} . e^{-\frac{1}{2}(y-\delta)^2} \quad \text{met} \quad \delta = \frac{\mu}{\sigma}\sqrt{2} \text{ en}$$

$$f_2(z) = \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad \text{voor } z \geqslant 0$$

van $\underline{y}$ en $\underline{z}$ worden afgeleid :

$$F(t) = P\left(\frac{\underline{y}}{\underline{z}} \leqslant t\right) = \int_0^\infty P(\underline{y} \leqslant zt)\, f_2(z)dz = \int_0^\infty F_1(zt) \quad f_2(z)dz.$$

Aan de voorwaarden voor differentiëren naar t onder het integraalteken is voldaan, dus

$$f(t) = \frac{dF(t)}{dt} = \int_0^\infty \frac{dF_1(zt)}{dt} f_2(z)\,dz = \int_0^\infty z f_1(zt).\, f_2(z)\,dz =$$

(6)

$$= \frac{1}{\pi} \int_0^\infty z.e^{-\frac{1}{2}\{z^2 + (zt - \delta)^2\}} dz.$$

We definiëren nu:

(7) $$\varphi(t) \overset{def}{=} - \frac{1}{\pi} \int_{-\infty}^0 z.\, e^{-\frac{1}{2}\{z^2 + (zt - \delta)^2\}} dz > 0 \text{ voor } -\infty < t < +\infty.$$

Dan geldt:

$$f(t) - \varphi(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} z.e^{-\frac{1}{2}\{z^2 + (zt - \delta)^2\}} dz =$$

$$= \frac{2}{\sqrt{2\pi}} \cdot \frac{\delta t}{(1+t^2)^{3/2}} \cdot e^{-\frac{1}{2} \cdot \frac{\delta^2}{1+t^2}} .$$

Dus geldt voor $T \geqslant 0$:

$$P(\underline{t} > T) = \int_T^\infty f(t)dt = \frac{2}{\sqrt{2\pi}} \int_T^\infty \frac{\delta t}{(1+t^2)^{3/2}} \cdot e^{-\frac{1}{2} \frac{\delta^2}{1+t^2}} dt + \int_T^\infty \varphi(t)dt$$

(8)

$$= \frac{2}{\sqrt{2\pi}} \cdot \int_0^{\frac{\delta}{\sqrt{1+T^2}}} e^{-\frac{1}{2}x^2} dx + \int_T^\infty \varphi(t)dt.$$

Daar $\varphi(t) > 0$ voor alle t, is

(9) $$\psi(T) \overset{def}{=} \int_T^\infty \varphi(t)dt$$

een monotoon dalende functie van T met $\lim\limits_{T \to \infty} \psi(T) = 0$.

Voorts is $P(\underline{t} > 0)$ exact bekend: .

(10) $$P(\underline{t} > 0) = P\left(\frac{\underline{y}}{\underline{z}} > 0\right) = P(\underline{y} > 0) = 1 - \emptyset(-\delta) = \emptyset(\delta),$$

waarin $\emptyset(\delta)$ de standaard-normale verdelingsfunctie voorstelt:

(11) $$\emptyset(\delta) \overset{def}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\delta} e^{-\frac{1}{2}x^2} dx.$$

Dus geldt volgens (8)

$$\psi(0) = P(\underline{t} > 0) - \frac{2}{\sqrt{2\pi}} \int_0^\delta e^{-\frac{1}{2}x^2} dx =$$

$$= \emptyset(\delta) - 2\left(\emptyset(\delta) - \tfrac{1}{2}\right) = 1 - \emptyset(\delta) = \emptyset(-\delta)$$

en wegens de monotonie van $\psi(T)$:

(12) $$0 < \psi(T) \leqslant \emptyset(-\delta) \quad \text{voor } 0 \leqslant T < \infty.$$

Uit (8) en (12) volgt dus voor $T \geqslant 0$

(13) $$2\,\emptyset\left(\frac{\delta}{\sqrt{1+T^2}}\right) - 1 < P(\underline{t} > T) \leqslant 2.\,\emptyset\left(\frac{\delta}{\sqrt{1+T^2}}\right) - 1 + \emptyset(-\delta).$$

Combinatie van (10) en (13) geeft voor $T \geqslant 0$

(14) $$2\,\phi\left(\frac{-\delta}{\sqrt{1+T^2}}\right) - 2\,\phi(-\delta) \leqslant P\,(0 \leqslant \underline{t} \leqslant T) < 2\,\phi\left(\frac{-\delta}{\sqrt{1+T^2}}\right) - \phi(-\delta).$$

Door hierin $\delta = \frac{\mu}{\sigma}\sqrt{2}$ en $T = \frac{2}{g_\alpha} \geqslant 0$ te kiezen en te substitueren in (5) wordt gevonden:

(15) $$\alpha + \phi\left(-\frac{\mu}{\sigma}\sqrt{2}\right) < 2\,\phi\left(\frac{-\frac{\mu}{\sigma}\sqrt{2}}{\sqrt{1+(^2/g_\alpha)^2}}\right) \leqslant \alpha + 2\,\phi\left(-\frac{\mu}{\sigma}\sqrt{2}\right)$$

Daar, volgens (1), $\frac{\mu}{\sigma}\sqrt{2} \geqslant 3\sqrt{2} = 4{,}24$ en dus $\phi\left(-\frac{\mu}{\sigma}\sqrt{2}\right) \leqslant 1{,}1.10^{-5}$, is $\phi\left(-\frac{\mu}{\sigma}\sqrt{2}\right)$ ten opzichte van α voor alle gebruikelijke waarden van α te verwaarlozen. g_α wordt dus gevonden uit

$$\frac{\frac{\mu}{\sigma}\sqrt{2}}{\sqrt{1+(^2/g_\alpha)^2}} = u_{\alpha/2} \quad , \text{ d.w.z.}$$

(3) $$g_\alpha = \frac{2\,u_{\alpha/2}}{\sqrt{2\left(\frac{\mu}{\sigma}\right)^2 - u^2_{\alpha/2}}} \; ,$$

waarin $u_{\alpha/2}$ gedefinieerd wordt door

(4) $$1 - \phi(u_{\alpha/2}) = \alpha/2.$$

Bij deze benadering wordt g_α overschat, of wel de bij g_α behorende waarde van α met een bedrag $\Delta\alpha$ onderschat, waarbij

$$1{,}1.10^{-5} \leqslant \phi\left(-\frac{\mu}{\sigma}\sqrt{2}\right) \leqslant \Delta\alpha \leqslant 2\,\phi\left(-\frac{\mu}{\sigma}\sqrt{2}\right) \leqslant 2{,}2.10^{-5}.$$

Nog afgezien van de geringe practische betekenis, lijkt het zinloos hiervoor een correctie op α aan te brengen, daar het feit, dat de werkelijke verdeling van $\underline{x}$ bij $x = 0$ en $x = 1$ afbreekt, een fout van dezelfde orde van grootte veroorzaakt.