STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

S 47 (M 17)

Transformatie op een homogene verdeling

MATHEMATISCH CENTRUM,
2de Boerhaavestr. 49,
Amsterdam-O.
Statistische Afdeling
S 47 (M 17)

Transformatie op een homogene verdeling.[1]

Ter toetsing van verschillende hypothesen kan men gebruik maken van een transformatie, waardoor een stochastische grootheid $\underline{x}$ met gegeven waarschijnlijkheidsverdeling, overgaat in een homogeen verdeelde grootheid $\underline{x}'$. Dit is een grootheid, waarvoor geldt:

$$(1) \quad P[\underline{x}' \leqq x'] = \begin{cases} 0 & \text{voor } x' < 0 \\ x' & \text{voor } 0 \leqq x' \leqq 1 \\ 1 & \text{voor } x' > 1 \end{cases}$$

dus met verdelingsdichtheid $h(x')$ (frequentiekromme) als in figuur 1

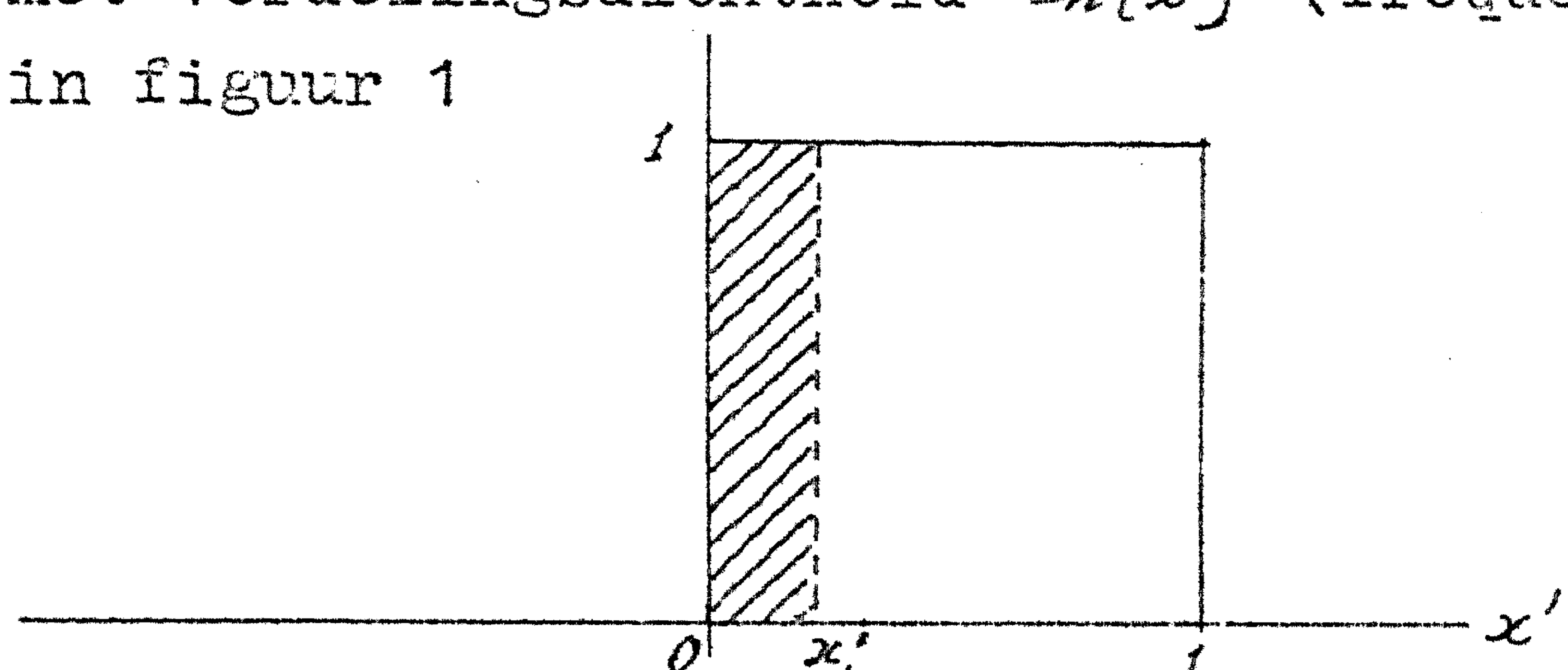

fig. 1

Indien nu de verdelingsdichtheid $f(x)$ van $\underline{x}$ gegeven wordt door de kromme van figuur 2:

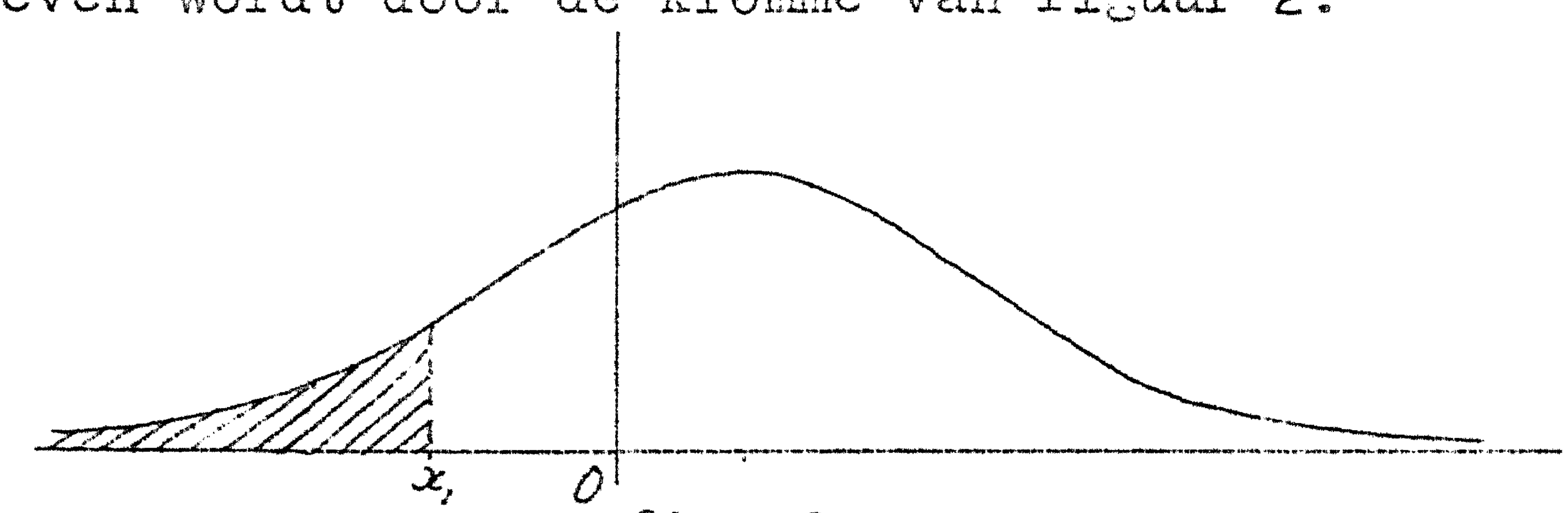

fig. 2

en we noemen de oppervlakte links van een waarde x_1 onder de kromme gelegen $F(x_1)$, dan bestaat deze transformatie op een homogene verdeling daaruit, dat x_1 vervangen wordt door dat punt x_1' (zie fig. 1), waarvoor geldt

$$(2) \qquad x_1' = F(x_1)$$

zodat links van x_1' dezelfde oppervlakte onder de kromme $h(x')$ ligt, als links van x_1 onder $f(x)$.

Men kan deze transformatie ook zeer duidelijk voorstellen door figuur 3, waarin F de verdelingsfunctie van $\underline{x}'$ en G die van $\underline{x}$ voorstelt.

---

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

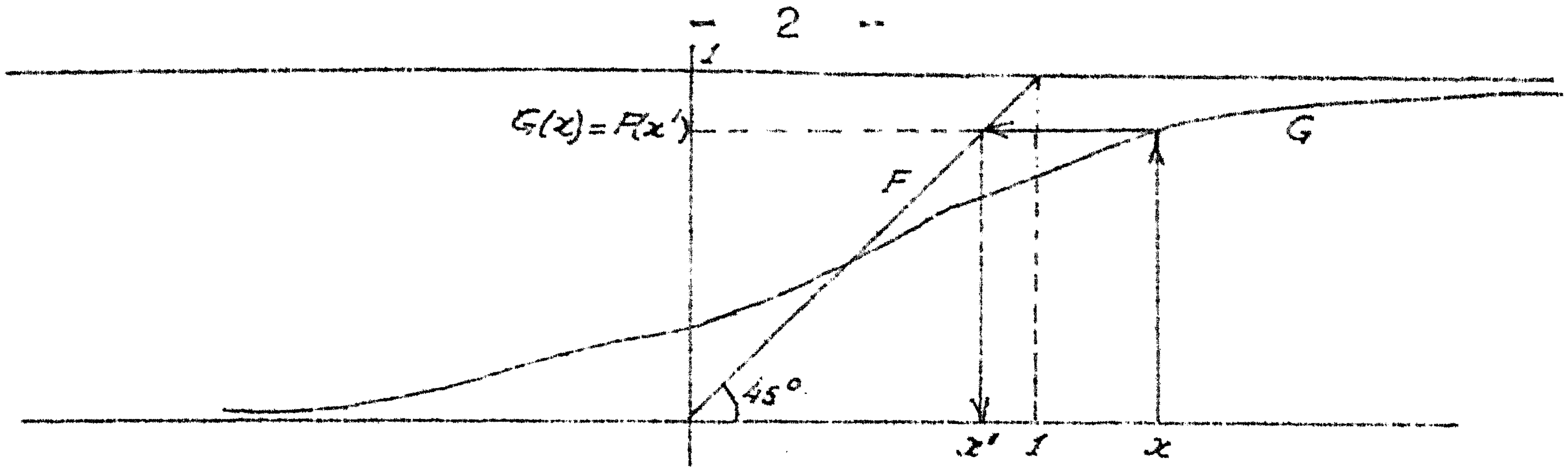

fig. 3

Voor een punt x is dus $G(x) = P[\underline{x} \leqq x]$. Zoeken wij nu het punt $\underline{x}'$, waarvoor $x' = F(x') = G(x)$ is, dan is dit het bij x behorende getransformeerde punt. De transformatie is in de figuur door een drietal pijlen aangegeven.

Willen wij nu de hypothese $\mathcal{H}_0$ toetsen, dat een aantal waarden $x_1, \ldots\ldots x_n$ een steekproef vormen uit een gegeven verdeling met verdelingsdichtheid $f(x)$, dan transformeren wij deze steekproef door de transformatie:

(3) $x_1' = F(x_1) , \ldots\ldots\ldots , x_n' = F(x_n)$

en toetsen in plaats van $\mathcal{H}_0$, de hypothese $\mathcal{H}_0^*$, dat $x_1', \ldots . x_n'$ een steekproef uit een homogene verdeling vormen.

Deze toets verloopt als volgt:
Vorm het product der waarden $x_1', \ldots . x_n'$ en zoek hiervan de natuurlijke logarithme op. Het getal

(4) $Q = -2\{\ln x_1' + \ln x_2' + \ldots\ldots\ldots \ln x_n'\}$ is

de bij de steekproef behorende waarde van een statistische grootheid $\underline{Q}$, die, zoals E.S.Pearson bewezen heeft, een χ^2-verdeling bezit met $2n$ vrijheidsgraden.

Is dus $Q \geqq Q_0$, waarin Q_0 een van n en de onbetrouwbaarheidsdrempel α afhankelijke grootheid is, die getabelleerd is, dan wordt $\mathcal{H}_0$ verworpen.

Literatuur:

E.S.Pearson, The Probability-Integral Transformation for testing goodness of fit and combining independent tests of significance. Biometrika 30 (1938) p.134-148.