STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

STATISTISCHE AFDELING

LEIDING: PROF. DR D. VAN DANTZIG
ADVISEUR VOOR STATISTISCHE CONSULTATIE: PROF. DR J. HEMELRIJK

Rapport S 250

De Gamma-verdeling

Overzicht betreffende eigenschappen
toetsings- en schattingsmethoden

door

Gerda Klerk-Grobben

april 1959

The Mathematical Centre at Amsterdam, founded the 11th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

Γ-verdelingen.

1. Inleiding.

1.1. Γ-functie. Voor de Γ-functie bestaan vele uitdrukkingen en formules (zie [26], hoofdstuk 12). De definitie van EULER is:

$$\Gamma(z) \stackrel{def}{=} \int_0^\infty e^{-t} t^{z-1} dt \quad ; \qquad Re\,(z) > 0.$$

De bekendste eigenschappen van deze functie zijn:

$$\Gamma(z+1) = z\,\Gamma(z) \quad ; \quad \Gamma(1) = 1 \quad ; \quad \Gamma(\tfrac{1}{2}) = \sqrt{\pi}\,.$$

$\Gamma(m) = (m-1)!$ voor m geheel en positief.

1.2. Γ-verdeling. De onvolledige Γ-functie ([26], pg.341) wordt gedefinieerd door:

$$\gamma(z,x) \stackrel{def}{=} \int_0^x e^{-t} t^{z-1} dt \quad ; \quad x > 0 \;, \; Re\,(z) > 0,$$

waaruit door vermenigvuldiging met $\frac{1}{\Gamma(z)}$ de verdelingsfunctie

$$G(x;z) \stackrel{def}{=} \frac{1}{\Gamma(z)} \int_0^x e^{-t} t^{z-1} dt$$

ontstaat. Hierin stelt z een parameter voor, die in dit rapport verder steeds met de letter γ aangegeven zal worden. De verdelingsfunctie $G(x;\gamma)$ kan nog gegeneraliseerd worden tot:

$$G(x;\alpha,\beta,\gamma) \stackrel{def}{=} \frac{1}{\beta^\gamma \Gamma(\gamma)} \int_\alpha^x e^{-\frac{t-\alpha}{\beta}} (t-\alpha)^{\gamma-1} dt \; ; \quad x > \alpha \,, \beta > 0 \,, \gamma > 0.$$

Alleen reële waarden van de parameters zullen worden beschouwd. $G(x;\alpha,\beta,\gamma)$ is de algemene gedaante van een verdeling van het type III in het systeem van PEARSON ([3] pg. 248 e.v.). Een verdeling met deze verdelingsfunctie zal een $\Gamma(\alpha,\beta,\gamma)$ - verdeling genoemd worden, de overeenkomstige verdelingsdichtheid wordt door $g(x;\alpha,\beta,\gamma)$ aangegeven. De parameter α geeft het beginpunt van de verdeling, β de schaaleenheid van de x-as, terwijl γ de vorm van de verdeling vastlegt.

1.3. Bijzondere gevallen. Door een speciale keuze van de parameters verkrijgt men enkele zeer bekende verdelingen, nl.

1.3.1. Gamma-verdeling met beginpunt bij 0, als $\alpha = 0$, dus met verdelingsdichtheid

$$g(x;0,\beta,\gamma) = \frac{1}{\beta^\gamma \Gamma(\gamma)} \, e^{-\frac{x}{\beta}} \, x^{\gamma-1}.$$

1.3.2. Exponentiële verdeling, als $\gamma = 1$, dus

$$g(x;\alpha,\beta,1) = \frac{1}{\beta} \, e^{-\frac{x-\alpha}{\beta}}.$$

1.3.3. χ^2_ν-verdeling, als $\alpha = 0$, $\beta = 2$ en $\gamma = \frac{\nu}{2}$, dus
$$g(x;0,2,\tfrac{\nu}{2}) = \frac{1}{2^{\frac{\nu}{2}}\,\Gamma(\frac{\nu}{2})}\, e^{-\frac{x}{2}}\, x^{\frac{\nu}{2}-1}.$$

2. Momenten.

2.1. Momenten genererende functie. De momenten genererende functie van een $\Gamma(\alpha,\beta,\gamma)$-verdeelde grootheid $\underline{x}$ wordt gegeven door
$$\mathcal{E}\, e^{\theta \underline{x}} = \frac{1}{\beta^\gamma \Gamma(\gamma)} \int_\alpha^\infty e^{\theta t}\, e^{-\frac{t-\alpha}{\beta}} (t-\alpha)^{\gamma-1}\, dt .$$
Substitutie van $u = t - \alpha$ geeft:

$$\mathcal{E}\, e^{\theta \underline{x}} = \frac{e^{\theta\alpha}}{\beta^\gamma \Gamma(\gamma)} \int_0^\infty e^{(\theta-\frac{1}{\beta})u}\, u^{\gamma-1}\, du = \frac{e^{\theta\alpha}}{\beta^\gamma \Gamma(\gamma)} \cdot \frac{\Gamma(\gamma)}{(\frac{1}{\beta}-\theta)^\gamma} = e^{\theta\alpha}(1-\beta\theta)^{-\gamma}.$$

De momenten om de oorsprong $(x = 0)$, welke uit de reeksontwikkeling naar θ van deze functie als coëfficiënt van $\frac{\theta^r}{r!}$ (voor het r^{de} moment, r = 1, 2, ...) gevonden worden, zijn vrij ingewikkelde uitdrukkingen in α, β en γ. Een eenvoudiger vorm bezitten de momenten ten opzichte van het beginpunt $(x = \alpha)$ van de verdeling. Deze volgen uit de reeksontwikkeling van
$$\mathcal{E}\, e^{\theta(\underline{x}-\alpha)} = (1-\theta\beta)^{-\gamma}.$$

Dit geeft:
$$\mathcal{E}(\underline{x}-\alpha)^r = \frac{\beta^r\, \Gamma(\gamma+r)}{\Gamma(\gamma)} .$$
De verwachting van $\underline{x}$ is dus
$$\mathcal{E}(\underline{x}) = \alpha + \beta\gamma ;$$
het 2e en 3e centrale moment (om $x = \alpha + \beta\gamma$):
$$\sigma^2(\underline{x}) = \gamma\beta^2 \quad \text{en} \quad \mathcal{E}(\underline{x}-\alpha-\beta\gamma)^3 = 2\gamma\beta^3 .$$
Voor de $\underline{\Gamma(0,\beta,\gamma)\text{-verdeling}}$ worden deze resp.
$$\mathcal{E}(\underline{x}) = \beta\gamma , \quad \sigma^2(\underline{x}) = \gamma\beta^2 \quad \text{en} \quad \mathcal{E}(\underline{x}-\beta\gamma)^3 = 2\gamma\beta^3 .$$
Voor de $\Gamma(\alpha,\beta,1)$ -verdeling (exponentiële verdeling):
$$\mathcal{E}(\underline{x}) = \alpha+\beta , \quad \sigma^2(\underline{x}) = \beta^2 \quad \text{en} \quad \mathcal{E}(\underline{x}-\alpha-\beta)^3 = 2\beta^3$$
en voor de χ^2_ν -verdeling $(\Gamma(0,2,\frac{\nu}{2}))$:
$$\mathcal{E}(\underline{x}) = \nu , \quad \sigma^2(\underline{x}) = 2\nu \quad \text{en} \quad \mathcal{E}(\underline{x}-\nu)^3 = 8\nu .$$

3. Gedaante van de verdelingsdichtheid; limietverdeling voor $\gamma \to \infty$.

3.1. Gedaante. De gedaante van de verdelingsdichtheid $g(x;\alpha,\beta,\gamma)$ hangt af van de waarde van de parameter γ ; het beginpunt en de schaal langs de x-as worden door α en β bepaald. Er kunnen 5 gevallen onderscheiden worden: $0 < \gamma < 1$; $\gamma = 1$; $1 < \gamma < 2$; $\gamma = 2$; $\gamma > 2$.

Voor elk van deze gevallen geldt dat de verdeling bij $x = \alpha$ begint, dat de verwachting $\mathcal{E}(\underline{x}) = \alpha + \beta\gamma$ is en de spreiding $\sigma(\underline{x}) = \beta\sqrt{\gamma}$. In figuren 1 t/m 5 zijn de 5 verschillende gevallen aangegeven, waarbij voor γ respectievelijk : $\gamma = 0{,}5$; $\gamma = 1$; $\gamma = 1{,}5$; $\gamma = 2$ en $\gamma = 4$ gekozen is.

3.1.1. (fig.1) $0 < \gamma < 1$

$$g(x;\alpha,\beta,\gamma) = \frac{1}{\beta^{\gamma}\Gamma(\gamma)}\, e^{-\frac{x-\alpha}{\beta}} (x-\alpha)^{\gamma-1},$$

$$\frac{d\,g(x)}{dx} = -\frac{1}{\beta^{\gamma}\Gamma(\gamma)}\,\frac{e^{-\frac{x-\alpha}{\beta}}}{(x-\alpha)^{1-\gamma}}\left(\frac{1}{\beta} + \frac{1-\gamma}{x-\alpha}\right),$$

$$\frac{d^2 g(x)}{dx^2} = \frac{1}{\beta^{\gamma}\Gamma(\gamma)}\,\frac{e^{-\frac{x-\alpha}{\beta}}}{(x-\alpha)^{1-\gamma}}\left(\frac{1}{\beta^2} + \frac{2(1-\gamma)}{\beta(x-\alpha)} + \frac{(1-\gamma)(2-\gamma)}{(x-\alpha)^2}\right).$$

Bij $x = \alpha$ is $g(x) = \infty$ en $\frac{d\,g(x)}{dx} = -\infty$, terwijl overal $\frac{d\,g(x)}{dx} < 0$ is, de verdelingsdichtheid is dus een monotoon dalende functie van x . De tweede en hogere afgeleiden bezitten geen nulpunten; de verdelingsdichtheid vertoont dus geen buigpunt.

3.1.2.(fig.2) $\gamma = 1$

$$g(x;\alpha,\beta,1) = \frac{1}{\beta}\, e^{-\frac{x-\alpha}{\beta}},$$

$$\frac{d\,g(x)}{dx} = -\frac{1}{\beta^2}\, e^{-\frac{x-\alpha}{\beta}},$$

In $x = \alpha$ is $g(x) = \frac{1}{\beta}$ en $\frac{d\,g(x)}{dx} = -\frac{1}{\beta^2}$. De tweede afgeleide blijft weer overal kleiner dan nul en de verdelingsdichtheid is dus ook nu nog monotoon dalend. Buigpunten zijn weer niet aanwezig.

3.1.3.(fig.3) $1 < \gamma < 2$

$$g(x;\alpha,\beta,\gamma) = \frac{1}{\beta^{\gamma}\Gamma(\gamma)}\, e^{-\frac{x-\alpha}{\beta}} (x-\alpha)^{\gamma-1},$$

$$\frac{d\,g(x)}{dx} = \frac{1}{\beta^{\gamma}\Gamma(\gamma)}\, e^{-\frac{x-\alpha}{\beta}} (x-\alpha)^{\gamma-1}\left(-\frac{1}{\beta} + \frac{\gamma-1}{x-\alpha}\right),$$

$$\frac{d^2 g(x)}{dx^2} = \frac{1}{\beta^{\gamma}\Gamma(\gamma)}\, e^{-\frac{x-\alpha}{\beta}} (x-\alpha)^{\gamma-1}\left(\frac{1}{\beta^2} - \frac{2(\gamma-1)}{\beta(x-\alpha)} - \frac{(\gamma-1)(2-\gamma)}{(x-\alpha)^2}\right).$$

In $x = \alpha$ is $g(x) = 0$ en $\frac{d\,g(x)}{dx} = +\infty$. De eerste afgeleide is nul, wanneer $\frac{1}{\beta} = \frac{\gamma - 1}{x - \alpha}$ of $x = \alpha + \beta(\gamma - 1)$, hier ligt dus de modus van de verdeling. De tweede afgeleide is 0 wanneer

$$(x-\alpha)^2 - 2\beta(\gamma-1)(x-\alpha) - \beta^2(\gamma-1)(2-\gamma) = 0$$

dus als

$$x = \alpha + \beta(\gamma - 1) \pm \beta\sqrt{\gamma - 1} ,$$

maar voor $1 < \gamma < 2$ is $\gamma - 1 < \sqrt{\gamma - 1}$ zodat alleen het + teken betekenis heeft; hier ligt dus het enige buigpunt van de verdeling.

3.1.4.(fig.4) $\gamma = 2$

$$g(x;\alpha,\beta,2) = \frac{1}{\beta^2}\, e^{-\frac{x-\alpha}{\beta}}(x-\alpha),$$

$$\frac{d\,g(x)}{dx} = \frac{1}{\beta^2}\, e^{-\frac{x-\alpha}{\beta}}\left(1 - \frac{x-\alpha}{\beta}\right),$$

$$\frac{d^2 g(x)}{dx^2} = \frac{1}{\beta^2}\, e^{-\frac{x-\alpha}{\beta}}\left(\frac{x-\alpha}{\beta^2} - \frac{2}{\beta}\right).$$

In $x = \alpha$ is $g(x) = 0$, $\frac{d\,g(x)}{dx} = \frac{1}{\beta^2}$. De eerste afgeleide is nul als $1 - \frac{x-\alpha}{\beta} = 0$, de modus van de verdeling ligt dus bij $x = \alpha + \beta$. De tweede afgeleide wordt nul als $\frac{x-\alpha}{\beta^2} - \frac{2}{\beta} = 0$, er is dus één buigpunt en wel bij $x = \alpha + 2\beta$ (in dit geval samenvallend met de verwachting van $\underline{x}$).

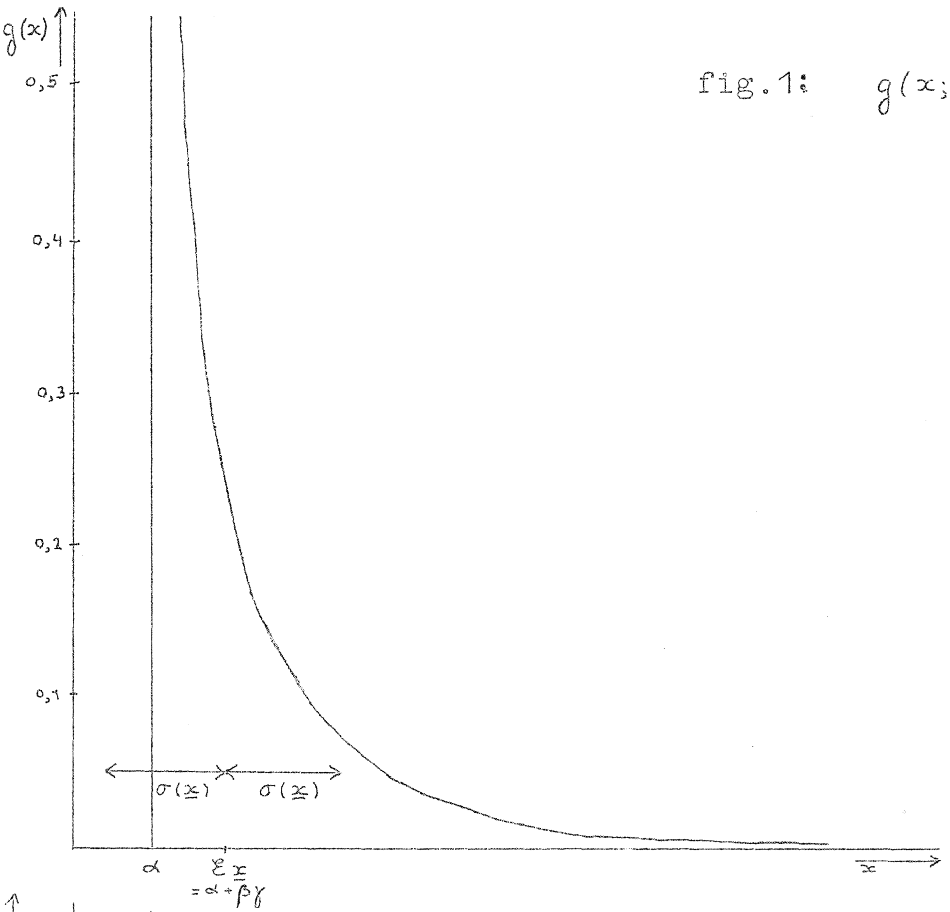


fig.1: $g(x;\alpha,\beta,\gamma)$ voor $0<\gamma<1$

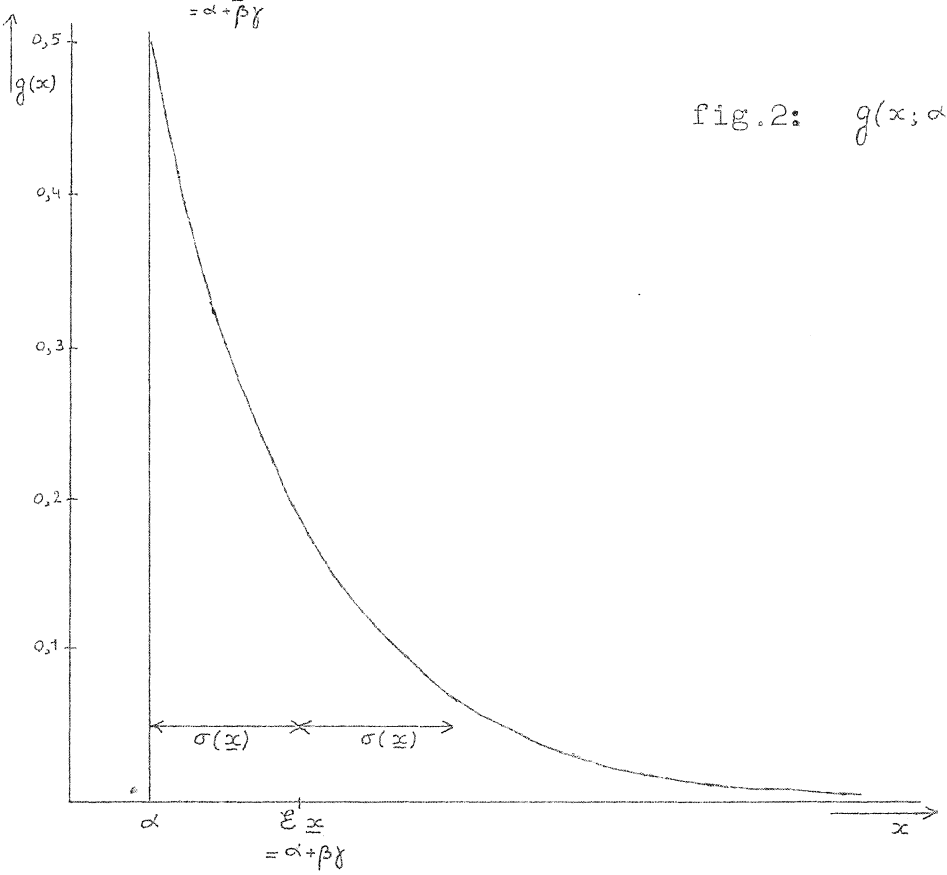


fig.2: $g(x;\alpha,\beta,\gamma)$ voor $\gamma=1$

fig.3: $g(x;\alpha,\beta,\gamma)$ voor $1<\gamma<2$

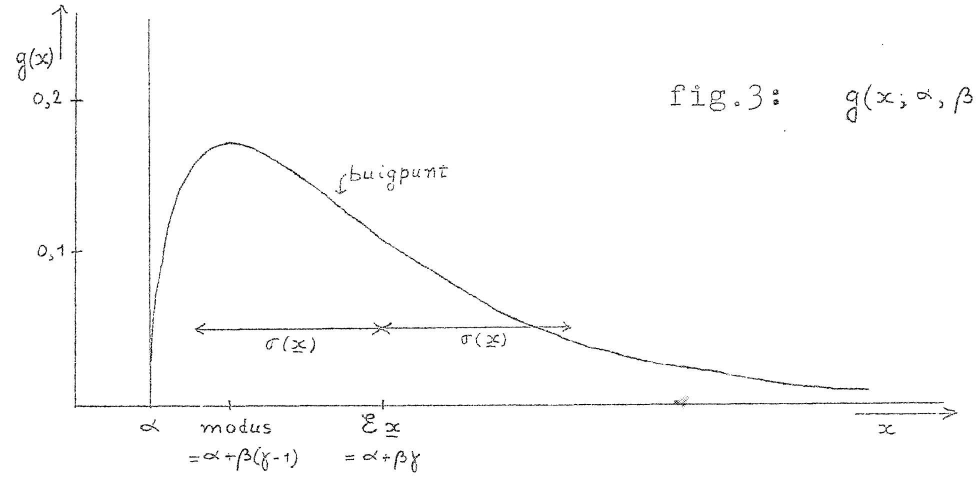


fig.4: $g(x;\alpha,\beta,\gamma)$ voor $\gamma = 2$

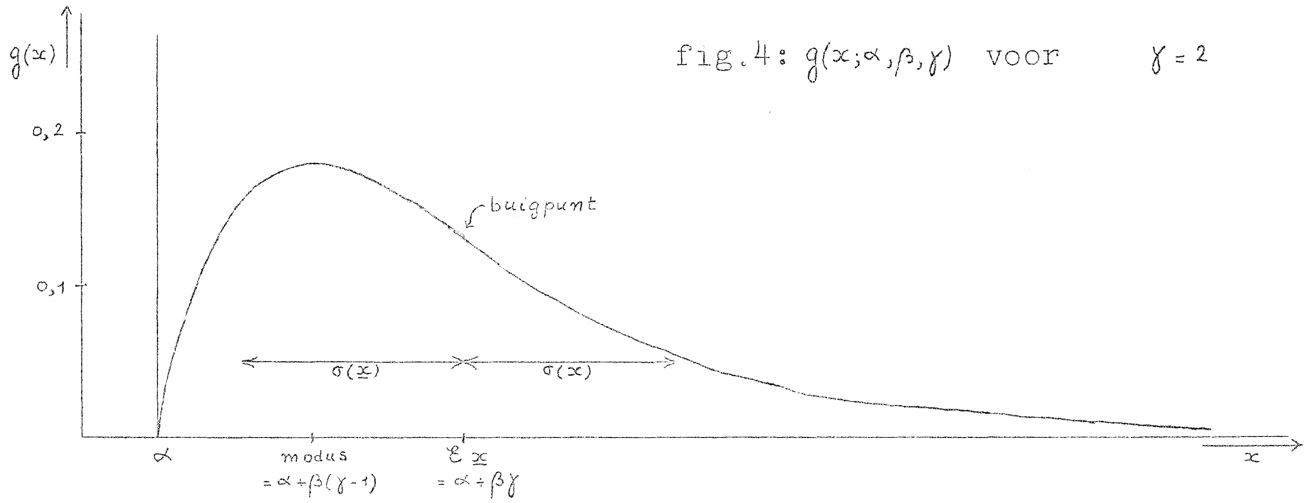
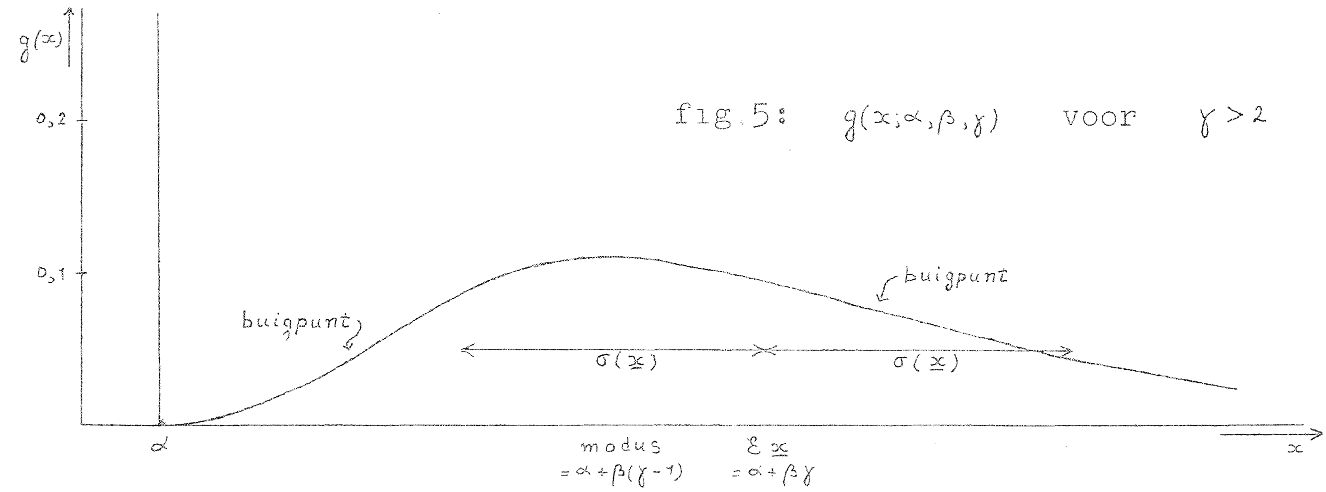


fig.5: $g(x;\alpha,\beta,\gamma)$ voor $\gamma > 2$



3.1.4. (fig.5), $\gamma > 2$

$$g(x;\alpha,\beta,\gamma) = \frac{1}{\beta^{\gamma}\Gamma(\gamma)} e^{-\frac{x-\alpha}{\beta}} (x-\alpha)^{\gamma-1},$$

$$\frac{dg(x)}{dx} = \frac{1}{\beta^{\gamma}\Gamma(\gamma)} e^{-\frac{x-\alpha}{\beta}} (x-\alpha)^{\gamma-2}\left(\gamma-1-\frac{x-\alpha}{\beta}\right),$$

$$\frac{d^2g(x)}{dx^2} = \frac{1}{\beta^{\gamma}\Gamma(\gamma)} e^{-\frac{x-\alpha}{\beta}} (x-\alpha)^{\gamma-2}\left(\frac{x-\alpha}{\beta^2} - \frac{2(\gamma-1)}{\beta} + \frac{(\gamma-1)(\gamma-2)}{x-\alpha}\right).$$

In $x = \alpha$ is $g(x) = 0$ en $\frac{dg(x)}{dx} = 0$. De eerste afgeleide wordt bovendien 0 als $\gamma - 1 - \frac{x-\alpha}{\beta} = 0$; de modus ligt dus bij $x = \alpha + \beta(\gamma-1)$. De tweede afgeleide wordt 0 als $(x-\alpha)^2 - 2\beta(\gamma-1)(x-\alpha) + (\gamma-1)(\gamma-2)\beta^2 = 0$. De buigpunten van de verdeling liggen dus bij $x = \alpha + \beta(\gamma-1) \pm \beta\sqrt{\gamma-1}$, dus op gelijke afstanden $\beta\sqrt{\gamma-1}$ ter weerszijden van de modus.
In de figuren zijn tevens de verwachting en de spreiding van $\underline{x}$ aangegeven.

3.2. <u>Limietverdeling voor $\gamma \to \infty$.</u> Indien $\underline{x}$ een $\Gamma(\alpha,\beta,\gamma)$ - verdeling bezit dan heeft $(\underline{x}-\alpha-\beta\gamma)/(\beta\sqrt{\gamma})$ voor $\gamma \to \infty$ een $N(0,1)$ -verdeling. Dit kan het eenvoudigst met behulp van de karakteristieke functie bewezen worden (Appendix 7.1). Voor grote waarden van γ heeft dus $\underline{x}$ bij benadering een normale verdeling met gemiddelde $\alpha+\beta\gamma$ en spreiding $\beta\sqrt{\gamma}$.
Worden alleen gehele waarden van γ beschouwd dan is een $\Gamma(\alpha,\beta,\gamma)$-verdeelde grootheid $\underline{x}$ op te vatten als de som van γ onafhankelijke grootheden met een $\Gamma(\frac{\alpha}{\gamma},\beta,1)$ -verdeling (zie par.5.2.).
De normaliteit van $\underline{x}$ voor $\gamma \to \infty$ volgt dan uit de centrale limietstelling.

4. Tabellen; Verband met Poisson-verdelingen.

4.1. Heeft $\underline{x}$ een $\Gamma(\alpha, \beta, \gamma)$ -verdeling dan heeft $\frac{2(\underline{x}-\alpha)}{\beta}$ een χ^2-verdeling met $\nu = 2\gamma$ vrijheidsgraden (zie par.1.3.3. en 5.1.). Tabellen van deze verdeling kunnen dus ook voor Γ-verdelingen gebruikt worden, eventueel met interpolatie indien 2γ niet geheel is. Een uitgebreide tabel van de χ^2-verdelingen is te vinden in de verzameling van E.S. PEARSON en H.O. HARTLEY [20]. In deze tabel van de χ^2_ν -verdelingsfuncties werden voor ν de waarden 1(1)30 (2) 70 gebruikt. Om G $(x; \alpha, \beta, \gamma)$ te bepalen moet in de tabel dus $\nu = 2\gamma$ en $\chi^2_\nu = \frac{2(x-\alpha)}{\beta}$ gebruikt worden:

$$G(x; \alpha, \beta, \gamma) = P[\underline{x} \leq x] = P\left[\frac{2(\underline{x}-\alpha)}{\beta} \leq \frac{2(x-\alpha)}{\beta}\right] = P\left[\underline{\chi}^2_{2\gamma} \leq \frac{2(x-\alpha)}{\beta}\right].$$

4.2. K. PEARSON [19] heeft uitgebreide tabellen samengesteld voor de verdelingsfunctie G$(x; \alpha, \beta, \gamma)$ als functie van $u = \frac{x-\alpha}{\beta\sqrt{\gamma}}$, waarbij als parameter niet γ, maar p = $\gamma - 1$ gebruikt is; de getabelleerde functie wordt met $I(u,p)$ aangegeven, dus:

$$G(x; \alpha, \beta, \gamma) = P[\underline{x} \leq x] = P[\underline{u} \leq u] =$$

$$= \frac{(\sqrt{p+1})^{p+1}}{\Gamma(p+1)} \int_0^u e^{-v\sqrt{p+1}} v^p \, dv =$$

$$= I(u,p) \quad \text{met} \quad u = \frac{x-\alpha}{\beta\sqrt{\gamma}} \quad \text{en} \quad p = \gamma - 1.$$

Voor p zijn de waarden:

p = -1,0(0,05)0,0(0,1)5,0(0,2)50,0

gebruikt, terwijl u met 0,1 opklimt.

4.3. L.R. SALVOSA [23] geeft zowel de verdelingsfunctie als de verdelingsdichtheid van de gestandaardiseerde $\Gamma(\alpha, \beta, \gamma)$-verdeling. In plaats van $\underline{u}$ wordt dus:

$$\underline{t} = \frac{\underline{x}-\alpha-\gamma\beta}{\beta\sqrt{\gamma}} \quad \text{met} \quad \mathcal{E}\underline{t} = 0 \quad \text{en} \quad \sigma^2(\underline{t}) = 1$$

gebruikt, waarbij nog $\gamma = \frac{4}{\alpha_3^2}$ gesteld wordt, zodat:

$$x = \alpha + \frac{4\beta}{\alpha_3^2}\left(1 + \frac{\alpha_3}{2} t\right) \quad \text{en} \quad \frac{dx}{dt} = \frac{2\beta}{\alpha_3} = \beta\sqrt{\gamma}.$$

De verdelingsfunctie wordt dan dus:

$$G(x;\alpha,\beta,\gamma) = P[\underline{x} \leq x] = P[\underline{t} \leq t] = c_0 \int_{\frac{-2}{\alpha_3}}^{t} e^{-\frac{2u}{\alpha_3}} \left(1 + \frac{\alpha_3}{2} u\right)^{\frac{4}{\alpha_3^2} - 1} du ,$$

met $$c_0 = \frac{\frac{\alpha_3}{2}\left(\frac{4}{\alpha_3^2}\right)^{\frac{4}{\alpha_3^2}} e^{-\frac{4}{\alpha_3^2}}}{\Gamma\left(\frac{4}{\alpha_3^2}\right)} ,$$

en de verdelingsdichtheid:

$$g(x;\alpha,\beta,\gamma) = f(t;\alpha_3) \cdot \frac{dt}{dx} = f(t;\alpha_3)/(\beta\sqrt{\gamma}) .$$

De functies $P[\underline{t} \leq t]$ en $f(t;\alpha_3)$ zijn dus getabelleerd met α_3 als parameter. Voor α_3 werden de waarden $\alpha_3 = 0(0,1)1,1$ gebruikt; dit komt voor γ overeen met de waarden:

$\gamma = \infty; 400; 100; 44,44; 25; 16; 11,11; 8,16; 6,25; 4,94; 4; 3,31.$

De variabele t klimt met 0,01 op.

4.4. <u>Poisson-verdeling.</u> De tabellen voor de verdelingsfunctie van de Γ-verdelingen kunnen ook bij het bepalen van overschrijdingskansen van POISSON-verdeelde grootheden gebruikt worden (zie [10] I pg. 122). Is nl. $\underline{n}$ POISSON-verdeeld dan geldt:

$$P[\underline{n} \geq n] = \sum_{i=n}^{\infty} \frac{\lambda^i}{i!} e^{-\lambda} = \frac{\lambda^n}{\Gamma(n)} \int_0^1 e^{-\lambda t} t^{n-1} dt = G(1; 0, \tfrac{1}{\lambda}, n) .$$

Men kan dus $P[\underline{n} \geq n]$ aflezen uit een χ^2-tabel, nl. als de kans dat een χ^2-verdeelde grootheid met $2n$ vrijheidsgraden groter is dan 2λ.

In de tabel van K. PEARSON is deze overschrijdingskans te vinden als $I(\frac{\lambda}{\sqrt{n}}, n-1)$, want $u = \frac{x-\alpha}{\beta\sqrt{\gamma}} = \frac{1-0}{\frac{1}{\lambda}\sqrt{n}} = \frac{\lambda}{\sqrt{n}}$ en $p = \gamma - 1 = n - 1$;

in de tabellen van Salvosa moet gezocht worden bij $\alpha_3 = \frac{2}{\sqrt{n}}$ en $t = \frac{\lambda - n}{\sqrt{n}}$.

Een ander verband dat bestaat tussen exponentiële verdelingen en POISSON-verdelingen zal in het volgende hoofdstuk behandeld worden (par. 5.5).

5. Eigenschappen van Γ-verdeelde grootheden.

In dit hoofdstuk worden een aantal eenvoudige eigenschappen van Γ-verdeelde grootheden vermeld. De meeste bewijzen zijn opgenomen in de appendix (hoofdstuk 7). Vele eigenschappen gelden alleen bij speciale waarden van de parameters.

5.1. Heeft $\underline{x}$ een $\Gamma(\alpha, \beta, \gamma)$-verdeling, dan bezit
$\underline{x}$ + a een $\Gamma(\alpha + a, \beta, \gamma)$ -verdeling en
b$\underline{x}$ een $\Gamma(b\alpha, b\beta, \gamma)$ -verdeling,
wanneer a en b constanten zijn en $b > 0$ (zie appendix 7.2.).
Gevolg: Heeft $\underline{x}$ een $\Gamma(\alpha, \beta, \gamma)$ -verdeling, dan heeft $2(\underline{x}-\alpha)/\beta$ een $\Gamma(0, 2, \gamma)$ – dus een $\chi^2_{2\gamma}$ -verdeling.

5.2. Zijn $\underline{x}_i$ (i=1, ..., n) onderling onafhankelijk en heeft $\underline{x}_i$ een $\Gamma(\alpha_i, \beta, \gamma_i)$-verdeling (i=1, ..., n) dan heeft $\sum_{i=1}^{n} \underline{x}_i$ een $\Gamma(\sum \alpha_i, \beta, \sum \gamma_i)$ -verdeling (appendix 7.3.).

5.3. Zijn $\underline{x}_1$ en $\underline{x}_2$ onafhankelijk verdeeld en bezit $\underline{x}_i$ een $\Gamma(0, \beta, \gamma_i)$ -verdeling (i=1 of 2), met gelijke β's dus, dan zijn ook $\underline{x}_1 + \underline{x}_2$ en $\underline{x}_1/\underline{x}_2$ onderling onafhankelijk verdeeld. Deze eigenschap is karakteristiek voor de Γ-verdeling, zoals de onafhankelijkheid van $\underline{x} + \underline{y}$ en $\underline{x} - \underline{y}$ karakteristiek is voor twee normaal verdeelde grootheden $\underline{x}$ en $\underline{y}$ met dezelfde variantie (zie [13] en [14]). Deze eigenschap kan als volgt tot meer onderling onafhankelijke grootheden gegeneraliseerd worden: Heeft x_i een $\Gamma(0, \beta, \gamma_i)$ -verdeling (i=1, ..., n) en zijn alle $\underline{x}_i$ onderling onafhankelijk, dan is $\sum_{i=1}^{n} \underline{x}_i$ onafhankelijk verdeeld van iedere homogene functie van de graad 0 in de $\underline{x}_i$, d.w.z. een functie $F(x_1, \ldots, x_n)$ die onafhankelijk is van de schaal langs de x-as, dus waarvoor

$$F(k x_1, k x_2, \ldots, k x_n) = F(x_1, x_2, \ldots, x_n)$$

(appendix 7.5 en [21]). In het bijzonder zijn ook $\underline{x}_1 + \underline{x}_2$ en $\underline{x}_1/(\underline{x}_1 + \underline{x}_2)$ onafhankelijk verdeeld. Voor $\Gamma(\alpha_i, \beta, \gamma_i)$ -verdeelde grootheden met $\alpha_i > 0$ gelden deze eigenschappen niet.

5.4. Zijn $\underline{x}_1$ en $\underline{x}_2$ verdeeld als onder 5.3. dan heeft $\underline{y} = \underline{x}_1/(\underline{x}_1 + \underline{x}_2)$ een β-verdeling (zie appendix 7.4 en [9], [12], [13], [14] en [21]), d.w.z. een verdeling met verdelingsdichtheid

$$\beta(y; \gamma_1, \gamma_2) = \frac{\Gamma(\gamma_1 + \gamma_2)}{\Gamma(\gamma_1)\Gamma(\gamma_2)} y^{\gamma_1 - 1}(1-y)^{\gamma_2 - 1}, \qquad (0 \leq y \leq 1).$$

De verdeling van $\underline{u} = \underline{x}_1/\underline{x}_2 = \underline{y}/(1-\underline{y})$ kan uit deze β-verdeling afgeleid worden; JAMBUNATHAN [9] spreekt van een β-verdeling van de 2e soort.

De relatie tussen exponentiële en POISSON-verdelingen wordt zonder bewijs vermeld:

5.5. Zijn $\underline{x}_1$, $\underline{x}_2$, ... onderling onafhankelijke $\Gamma(0,\beta,1)$-verdeelde grootheden dan bezit het aantal $\underline{n}_T$, dat vastgelegd wordt door de eisen:

$$\sum_{i=1}^{n_T} \underline{x}_i \leq T \quad \text{en} \quad \sum_{i=1}^{n_T+1} \underline{x}_i > T$$

een POISSON-verdeling met parameter $\frac{T}{\beta}$:

$$P[\underline{n}_T = n] = e^{-\frac{T}{\beta}} \left(\frac{T}{\beta}\right)^n / n!\,.$$

Ook geldt het omgekeerde: Zijn $\underline{x}_1$, $\underline{x}_2$, ... onderling onafhankelijke stochastische grootheden met eenzelfde niet-ontaarde verdeling en bezit het aantal $\underline{n}_T$, dat door bovenstaande eisen is vastgelegd, een POISSON-verdeling met parameter λ dan zijn $\underline{x}_1$, $\underline{x}_2$, ... afkomstig uit een $\Gamma(0,\frac{T}{\lambda},1)$ -verdeling (zie [17]). Deze stellingen kunnen nog gegeneraliseerd worden tot eenzelfde verband tussen een $\Gamma(0,\beta,2)$ -verdeling en een gegeneraliseerde POISSON-verdeling [7].

Dit hoofdstuk wordt besloten met een reeks eigenschappen van exponentiële verdelingen, welke PITMAN in [21] geeft. De waarnemingen $\underline{x}_i$ bezitten dus een $\Gamma(\alpha,\beta,1)$ -verdeling. De kleinste waarde uit de steekproef $\underline{x}_1$, ..., $\underline{x}_n$ zal met $\underline{x}_0$ worden aangegeven:

$$\underline{x}_0 \stackrel{\text{def}}{=} \min(\underline{x}_1, \ldots, \underline{x}_n)\,.$$

5.6. Voor exponentieel verdeelde grootheden geldt:

1. $\underline{x}_0$ heeft een $\Gamma(\alpha,\frac{\beta}{n},1)$ -verdeling,
2. $\sum_{i=1}^{n}(\underline{x}_i-\underline{x}_0)$ heeft een $\Gamma(0,\beta,n-1)$ -verdeling en
3. $\underline{x}_0$ en $\sum_{i=1}^{n}(\underline{x}_i-\underline{x}_0)$ zijn onderling onafhankelijk (zie appendix 7.6.).

5.7. Is $T(x_1, \ldots, x_n)$ een functie van $x_1, \ldots, x_n$ die onafhankelijk is van de keuze van de oorsprong (dus van de waarde van α, zoals $\sum_{i=1}^{n}(\underline{x}_i-\underline{x}_0)$ maar ook bijv. $\underline{x}_n-\underline{x}_0$) dan zijn $\underline{T} = T(\underline{x}_1, \ldots, \underline{x}_n)$ en $\underline{x}_0$ onderling onafhankelijk (appendix 7.7.).

5.8. Is $G(x_1, \ldots, x_n)$ een homogene functie van de graad 0 , dus waarvoor geldt:

$$G(k x_1, \ldots, k x_n) = G(x_1, \ldots, x_n)\,,$$

die bovendien onafhankelijk is van de keuze van de oorsprong,dan zijn $\underline{G}$, $\underline{x}_0$ en $\sum_{i=1}^{n}(\underline{x}_i-\underline{x}_0)$ onderling onafhankelijk verdeeld (appendix 7.8.).

5.9. Uit het feit dat $\underline{x}_0, \sum_{i=1}^{n}(\underline{x}_i-\underline{x}_0)$ en $\underline{G}$ (zie par. 5.8.) onderling onafhankelijk zijn volgt dat $\underline{G}$ onafhankelijk verdeeld is van een willekeurige functie van $\underline{x}_0$ en $\sum_{i=1}^{n}(\underline{x}_i-\underline{x}_0)$.

6. Schattingen, toetsen en betrouwbaarheidsintervallen.

In de litteratuur is weinig bekend over schattingen en toetsen bij een algemene $\Gamma(\alpha,\beta,\gamma)$-verdeling. Zijn echter één of twee parameters bekend dan wordt de situatie iets gunstiger, terwijl voor de exponentiële verdeling $(\gamma = 1)$ een hele reeks toetsen en betrouwbaarheidsintervallen ter beschikking staan.

In het volgende wordt bij een toets de toetsingsgrootheid steeds met de letter T aangegeven voorzien van een index die gelijk is aan het paragraafnummer. Is $\underline{T}_i$ de toetsingsgrootheid en $f(T_i)$ de bijbehorende verdelingsdichtheid onder H_0 dan wordt met $T_i(\varepsilon)$ die waarde bedoeld, waarvoor:

$$\int_{-\infty}^{T_i(\varepsilon)} f(T_i)\, dT_i = \varepsilon . \qquad (6;1)$$

Tenzij anders aangegeven, wordt steeds ondersteld dat $\underline{x}_1, \ldots, \underline{x}_n$ onderling onafhankelijke waarnemingen zijn uit eenzelfde Γ-verdeling.

6.1. Om uit een steekproef $\underline{x}_1, \ldots, \underline{x}_n$ uit een $\Gamma(\alpha,\beta,\gamma)$-verdeling de drie onbekende parameters te schatten kunnen de schattingen voor de momenten gebruikt worden:

$\underline{\bar{x}} = \sum_{i=1}^{n} \underline{x}_i/n$ als schatting voor $\mathcal{E}\,\underline{x} = \alpha + \beta\gamma$,

$\underline{s}^2 = \sum_{i=1}^{n}(\underline{x}_i - \underline{\bar{x}})^2/n$ als schatting voor $\sigma^2(\underline{x}) = \gamma\beta^2$,

$\underline{m}_3 = \sum_{i=1}^{n}(\underline{x}_i - \underline{\bar{x}})^3/n$ als schatting voor $\mathcal{E}(\underline{x} - \mathcal{E}\,\underline{x})^3 = 2\gamma\beta^3$.

Voor α, β en γ volgen dan respectievelijk de schattingen:

$$\underline{a}_1 = \underline{\bar{x}} - \frac{2(\underline{s}^2)^2}{\underline{m}_3},$$

$$\underline{b}_1 = \frac{\underline{m}_3}{2\underline{s}^2},$$

en

$$\underline{c}_1 = \frac{4(\underline{s}^2)^3}{\underline{m}_3^2}.$$

GREENWOOD en DURAND [8] geven een iteratiemethode voor het bepalen van meest aannemelijke schattingen, waarbij ze gebruik maken van 't bepalen van aannemelijkste schattingen bij gegeven α (zie par. 6.2.).

6.2. Is de parameter α bekend, dan kan zonder verlies van algemeenheid ondersteld worden dat alle $\underline{x}_i$ een $\Gamma(o,\beta,\gamma)$ -verdeling bezitten. Uit het gemiddelde en de spreiding van de steekproef volgen dan

(6.2;1) $\quad \underline{b}_2 = \frac{\underline{s}^2}{\underline{\bar{x}}} \quad$ en $\quad \underline{c}_2 = \frac{\underline{\bar{x}}^2}{\underline{s}^2}$

als schattingen voor β en γ.

In plaats van deze momentenmethode kan ook de methode der meest aannemelijke schattingen gebruikt worden (zie appendix 7.9). Hierbij blijkt dan dat voor de aannemelijkste schattingen $\hat{\underline{b}}$ en $\hat{\underline{c}}$ geldt:

$$\ln \hat{\underline{c}} - \frac{\Gamma(\hat{\underline{c}})}{\Gamma'(\hat{\underline{c}})} = \ln \frac{\sum_{i=1}^{n} \underline{x}_i}{n} - \frac{\sum_{i=1}^{n} \ln \underline{x}_i}{n}$$

en

$$\hat{\underline{b}} = \frac{\underline{\bar{x}}}{\hat{\underline{c}}} .$$

Uit een tabel van de functie $\ln t - \Gamma(t)/\Gamma'(t)$ is dan, bij de waarde van $\ln(\sum_{i=1}^{n} \underline{x}_i/n) - (\sum_{i=1}^{n} \ln \underline{x}_i)/n$ verkregen met behulp van de steekproef, de bijbehorende waarde van t dus van $\hat{\underline{c}}$ op te zoeken (zie GREENWOOD en DURAND [8]). Als benadering voor deze aannemelijkste schattingen kan men gebruiken:

(6.2;2) $\underline{c}^* = \dfrac{1}{2(\ln \underline{\bar{x}} - \frac{\sum_{i=1}^{n} \ln \underline{x}_i}{n})} \quad$ en $\quad \underline{b}^* = \dfrac{\underline{\bar{x}}}{\underline{c}^*}$.

De nauwkeurigheid van de drie stellen schattingen uit deze paragraaf zijn met behulp van een aantal steekproef experimenten onderzocht (zie appendix 7.10). De resultaten doen vermoeden dat bij alle drie de variantie van de schatting vrij groot is, althans bij $\gamma \geq 4$. Dit is in overeenstemming met de benaderingsformule voor $\sigma^2(\hat{\underline{c}})$ van GREENWOOD en DURAND [8].

6.3. Bezitten $\underline{x}_1$, ..., $\underline{x}_n$ een $\Gamma(o,\beta,\gamma)$ -verdeling met <u>bekende γ</u> dan is:

$$\underline{b}_3 = \frac{\underline{\bar{x}}}{\gamma}$$

een zuivere aannemelijkste <u>schatting</u> voor β (volgt uit de vergelijkingen in appendix 7.9.) en

$$\frac{1}{\underline{b}'} = \frac{n\gamma - 1}{n\underline{\bar{x}}}$$

een zuivere <u>schatting</u> voor $\frac{1}{\beta}$ (zie appendix 7.11. en [10] II en [22]). Uit eigenschap 5.2. volgt dat $\sum_{i=1}^{n} \underline{x}_i$ een $\Gamma(o,\beta,n\gamma)$ -verdeling bezit of wel $\frac{2\sum_{i=1}^{n} \underline{x}_i}{\beta}$ een $\Gamma(o,2,n\gamma)$- dus een $\chi^2_{2n\gamma}$-verdeling. Met behulp hiervan kan een <u>toets</u> voor de hypothese H_0: $\beta = \beta_0$ opgesteld worden. Toetsingsgrootheid hierbij is dan:

$$\underline{T}_3 = \frac{2\sum_{i=1}^{n} \underline{x}_i}{\beta_0} .$$

welke dus onder H_0 een $\Gamma(0, 2, n\gamma)$ -verdeling heeft. Bij een gevonden waarde van T_3 zijn de overschrijdingskansen k_ℓ (links) en k_r (rechts):

$$k_\ell = \int_0^{T_3} \frac{1}{2^{n\gamma}\Gamma(n\gamma)} e^{-\frac{t}{2}} t^{n\gamma-1} dt = G(T_3; 0; 2; n\gamma) \quad \text{en} \quad k_r = 1 - k_\ell .$$

De tweezijdige overschrijdingskans kan dan gedefinieerd worden als

$$k = \min(k_r \text{ en } k_\ell),$$

(zie [10] II en ook [2] voor het onderscheidingsvermogen van deze toets). Is $T_3(\varepsilon)$ gedefinieerd als in 't begin van dit hoofdstuk dan is het interval:

$$\left(2\sum_{i=1}^{n} x_i / T_3(1-\varepsilon) \, ; \, 2\sum_{i=1}^{n} x_i / T_3(\varepsilon)\right)$$

een tweezijdig begrensd <u>betrouwbaarheidsinterval</u> voor β met onbetrouwbaarheid 2ε.

<u>Opmerking:</u> In plaats van met dezelfde γ, mogen de grootheden $\underline{x}_i$ ook uit verdelingen met verschillende γ_i afkomstig zijn. In de formules moet dan overal $n\gamma$ door $\sum_{i=1}^{n}\gamma_i$ vervangen worden.

6.4. Bezitten alle $\underline{x}_i$ een $\Gamma(\alpha, \beta, \gamma)$-verdeling, waarin <u>$\beta$ en γ bekend</u> zijn dan heeft $\sum_{i=1}^{n} \underline{x}_i$ een $\Gamma(n\alpha, \beta, n\gamma)$ -verdeling en dus $\frac{2(\sum_{i=1}^{n} x_i - n\alpha)}{\beta}$ een $\Gamma(0, 2, n\gamma)$- of een $\chi^2_{2n\gamma}$-verdeling. Een zuivere <u>schatting</u> voor α is :

$$a_4 = \bar{x} - \beta\gamma .$$

Om de hypothese $H_0: \alpha = \alpha_0$ te <u>toetsen</u> kan als toetsingsgrootheid

$$T_4 = 2\left(\sum_{i=1}^{n} x_i - n\alpha_0\right)/\beta$$

gebruikt worden, welke onder H_0 dus een $\Gamma(0, 2, n\gamma)$ - of een $\chi^2_{2n\gamma}$-verdeling bezit. De overschrijdingskansen van een gevonden waarde voor T_4 kunnen weer als in par. 6.3. bepaald worden. Een <u>betrouwbaarheidsinterval</u> voor α met onbetrouwbaarheid 2ε is:

$$\left(\bar{x} - \beta T_4(1-\varepsilon)/(2n) \, ; \, \bar{x} - \beta T_4(\varepsilon)/(2n)\right),$$

waarin $T_4(1-\varepsilon)$ en $T_4(\varepsilon)$ weer door (6;1) gedefinieerd zijn.

<u>Opmerking:</u> Ook hier mogen de grootheden $\underline{x}_i$ uit verdelingen met verschillende γ_i afkomstig zijn.

De volgende paragrafen hebben alle betrekking op exponentiële verdelingen, waarbij dus $\gamma = 1$ is. Zoals in hoofdstuk 5 wordt ook nu $\underline{x}_0$ gedefinieerd als de kleinste waarde uit een steekproef $\underline{x}_1, \ldots, \underline{x}_n$.

6.5. De waarnemingen $\underline{x}_i$ zijn nu afkomstig uit een $\Gamma(\alpha, \beta, 1)$-verdeling. Volgens par. 5.6. heeft dan $\sum_{i=1}^{n} (\underline{x}_i - \underline{x}_0)$ een $\Gamma(0, \beta, n-1)$ -verdeling, dus met

$$\mathcal{E} \sum_{i=1}^{n} (\underline{x}_i - \underline{x}_0) = (n-1)\beta .$$

De schatting

$$\underline{b}_5 = \sum_{i=1}^{n} (\underline{x}_i - \underline{x}_0)/(n-1)$$

voor β is dan ook zuiver en bezit bovendien een minimale variantie (zie [5] en [6]). Om de hypothese H_0: $\beta = \beta_0$ te toetsen kan

$$\underline{T}_5 = \frac{2(n-1)\,\underline{b}_5}{\beta_0} = \frac{2\sum_{i=1}^{n} (\underline{x}_i - \underline{x}_0)}{\beta_0}$$

als toetsingsgrootheid gebruikt worden, welke onder H_0 een $\Gamma(0, 2, n-1)$ - dus een $\chi^2_{2(n-1)}$-verdeling bezit. Een tweezijdig begrensd betrouwbaarheidsinterval met onbetrouwbaarheid 2ε dat hieruit volgt is :

$$\left(\frac{2\sum_{i=1}^{n} (\underline{x}_i - \underline{x}_0)}{\chi^2_{2(n-1)}(1-\varepsilon)} \; ; \; \frac{2\sum_{i=1}^{n} (\underline{x}_i - \underline{x}_0)}{\chi^2_{2(n-1)}(\varepsilon)} \right) ,$$

waarin $\chi^2_{2(n-1)}(1-\varepsilon)$ en $\chi^2_{2(n-1)}(\varepsilon)$ resp. rechter en linker kritieke waarde uit de $\chi^2_{2(n-1)}$-verdeling zijn, analoog aan (6;1) gedefinieerd.

6.6. EPSTEIN en SOBEL [6] hebben het in de vorige paragraaf behandelde geval algemener behandeld. Zij onderstellen dat de waarnemingen $\underline{x}_1$, ..., $\underline{x}_n$ afkomstig zijn uit k verschillende exponentiële verdelingen, die echter alle dezelfde (onbekende) parameter β bezitten. Voor de α_i $(i = 1, \ldots, k)$ worden drie mogelijkheden onderscheiden:

1) De α_i zijn bekend en ongelijk: de schatting voor β is nu :

$$\underline{b}_{6,1} = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (\underline{x}_{ji} - \alpha_j)/n \qquad \text{met} \qquad \sum_{j=1}^{k} n_j = n ,$$

terwijl de toetsingsgrootheid

$$\underline{T}_{6,1} = 2n\,\underline{b}_{6,1}/\beta_0$$

een χ^2_{2n} -verdeling bezit onder H_0. $(\beta = \beta_0)$.

2) De α_i zijn onbekend maar alle gelijk, dit is het geval van paragraaf 6.5.

3) De α_i zijn onbekend en ongelijk; de schatting voor β wordt:

$$\underline{b}_{6,3} = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (\underline{x}_{ji} - \underline{x}_{j0}) / (n-k) \qquad \text{met} \qquad \underline{x}_{j0} \overset{def}{=\!=} min(\underline{x}_{j1}, \ldots, \underline{x}_{jn_j}),$$

en de toetsingsgrootheid

$$\underline{T}_{6,3} = 2(n-k)\, \underline{b}_{6,3} / \beta_0$$

heeft een $\chi^2_{2(n-k)}$-verdeling onder H_0 $(\beta = \beta_0)$.

6.7. Zijn de waarnemingen $\underline{x}_i$ als in par. 6.5. dan is

$$\underline{a}_7 = \underline{x}_0 - \frac{\sum_{i=1}^{n} (\underline{x}_i - \underline{x}_0)}{n(n-1)}$$

een zuivere schatting voor α met minimale variantie (zie [6]). Een betrouwbaarheidsinterval voor α en een toets voor een hypothese H_0: $\alpha = \alpha_0$ volgen uit de eigenschappen in par.5.6. en par. 5.4.

Immers nu heeft

$$\underline{T}_7 = \frac{n(\underline{x}_0 - \alpha_0)}{\sum_{i=1}^{n} (\underline{x}_i - \underline{x}_0) + n(\underline{x}_0 - \alpha_0)} = \frac{n(\underline{x}_0 - \alpha_0)}{\sum_{i=1}^{n} (\underline{x}_i - \alpha_0)}$$

onder H_0 een $\beta(1, n-1)$ -verdeling, dus met verdelingsdichtheid:

$$\beta(T_7\, ; 1, n-1) = (n-1)(1-T_7)^{n-2}, \qquad 0 \le T_7 \le 1.$$

De overschrijdingskansen bij een gevonden waarde T_7 zijn:

$$k_\ell = \int_0^{T_7} (n-1)(1-u)^{n-2}\, du = 1-(1-T_7)^{n-1}, \qquad k_r = (1-T_7)^{n-1}.$$

Toepassing van (6;1) geeft voor de kritieke waarde $T_7(\varepsilon)$:

$$\varepsilon = \int_0^{T_7(\varepsilon)} (n-1)(1-u)^{n-2}\, du = 1-(1-T_7(\varepsilon))^{n-1}$$

of $T_7(\varepsilon) = 1-(1-\varepsilon)^{\frac{1}{n-1}}$ en analoog $T_7(1-\varepsilon) = 1-\varepsilon^{\frac{1}{n-1}}$. Als betrouwbaarheidsinterval met onbetrouwbaarheid ε is hieruit te vinden :

$$\left\{ (\underline{x}_0 - \bar{\underline{x}}\, T_7(\varepsilon)) / (1 - T_7(\varepsilon))\, ; \underline{x}_0 \right\} \quad \text{(zie appendix 7.12).}$$

Hoewel het interval tweezijdig begrensd is, is de onbetrouwbaarheid toch ε en niet 2ε, daar:

$$P\left[\alpha > \underline{x}_0 \right] = 0.$$

6.8. Bij bepaalde typen proeven, zoals bijv. levensduur proeven (zie [5] en [6]), worden de waarnemingen in volgorde van opklimmende grootte verkregen. In een dergelijk geval kan de proef, ter bekorting, beëindigd worden voordat alle waarnemingen verricht zijn. Stel de waar te nemen grootheid heeft een $\Gamma(0,\beta,1)$ -verdeling. De waarnemingen zijn $\underline{x}_{(1)}, \ldots, \underline{x}_{(r)}$ (indices tussen haakjes worden vaak gebruikt om aan te geven dat de waarnemingen naar grootte gerangschikt staan), met nog als extra informatie dat de $n-r$ overige waarnemingen groter dan $\underline{x}_{(r)}$ zouden zijn. EPSTEIN en SOBEL([5] en [6]) bewijzen dat :

$$\underline{b}_{r,n} = \frac{1}{r}\left[\sum_{i=1}^{r} \underline{x}_{(i)} + (n-r)\,\underline{x}_{(r)}\right]$$

een meest aannemelijke schatting voor β is. Voor een toets voor een hypothese $H_0 : \beta = \beta_0$ kan als toetsingsgrootheid :

$$\underline{T}_8 = \frac{2r\,\underline{b}_{r,n}}{\beta_0}$$

gebruikt worden, welke onder H_0 een χ^2_{2r} -verdeling bezit. De kritieke waarden $\chi^2_{2r}(\varepsilon)$ en $\chi^2_{2r}(1-\varepsilon)$ uit deze verdeling geven:

$$\left(\frac{2r\,\underline{b}_{r,n}}{\chi^2_{2r}(1-\varepsilon)}\ ;\ \frac{2r\,\underline{b}_{r,n}}{\chi^2_{2r}(\varepsilon)}\right)$$

als tweezijdig begrensd betrouwbaarheidsinterval voor β, met onbetrouwbaarheid 2ε.

Dit hoofdstuk wordt besloten met enkele toetsen voor het vergelijken van twee of meer exponentiële verdelingen.

6.9. Hebben $\underline{x}_1, \ldots, \underline{x}_n$ een $\Gamma(\alpha,\beta,1)$ -verdeling en $\underline{x}'_1, \ldots, \underline{x}'_n$ een $\Gamma(\alpha',\beta,1)$ -verdeling, waarbij β bekend is, dan heeft volgens par. 5.6. $\underline{x}_0 = \min(\underline{x}_1, \ldots, \underline{x}_n)$ een $\Gamma(\alpha, \frac{\beta}{n}, 1)$ -verdeling en $\underline{x}'_0 = \min(\underline{x}'_1, \ldots, \underline{x}'_n)$ een $\Gamma(\alpha', \frac{\beta}{n}, 1)$ -verdeling. Dit kan gebruikt worden om de hypothese $H_0 : \alpha = \alpha' + \delta_0$ te toetsen. Toetsingsgrootheid is dan:

$$\underline{T}_9 = n(\underline{x}_0 - \underline{x}'_0 - \delta_0)/\beta,$$

welke volgens KULLBACK [12] een verdeling met verdelingsdichtheid $\frac{1}{2}e^{-|T_9|}$ $(-\infty < T_9 < \infty)$ heeft. De tweezijdige overschrijdingskans is dus:

$$2P\left[\underline{T}_g \geq |T_g|\right] = \int_{|T_g|}^{\infty} e^{-t}\,dt = e^{-|T_g|} .$$

Is $T_g(\varepsilon)$ de waarde waarvoor $e^{-T_g(\varepsilon)} = \varepsilon$, dus $T_g(\varepsilon) = -\ln\varepsilon$, dan is

$$\left(\underline{x}_o - \underline{x}_o' - \beta T_g(\varepsilon)/n \; ; \; \underline{x}_o - \underline{x}_o' + \beta T_g(\varepsilon)/n\right)$$

een <u>betrouwbaarheidsinterval</u> voor δ met onbetrouwbaarheid ε .
Bezitten $\underline{x}_i$ en $\underline{x}_i'$ exponentiële verdelingen met verschillende maar wel <u>bekende</u> β en β' , dan kan 't bovenstaande niet zo worden toegepast mits nu ook de aantallen waarnemingen n en n' verschillend zijn en wel zodanig, dat $\frac{\beta}{n} = \frac{\beta'}{n'}$.

6.10. Is $\underline{x}_1, \ldots, \underline{x}_n$ een steekproef uit een $\Gamma(\alpha, \beta, 1)$ -verdeling en $\underline{x}_1', \ldots, \underline{x}_n'$, uit een $\Gamma(\alpha', \beta', 1)$-verdeling dan heeft volgens par. 5.6 $\sum_{i=1}^{n} (\underline{x}_i - \underline{x}_0)/\beta$ een $\Gamma(0, 1, n-1)$-verdeling en $\sum_{i=1}^{n'} (\underline{x}_i' - \underline{x}_0')/\beta'$ een $\Gamma(0, 1, n'-1)$-verdeling. Om de hypothese $H_o: \beta = \kappa_o . \beta'$ te toetsen kan de toetsingsgrootheid

$$\underline{T}_{10} = \frac{\sum_{i=1}^{n} (\underline{x}_i - \underline{x}_o)}{\kappa_o \sum_{i=1}^{n'} (\underline{x}_i' - \underline{x}_o')}$$

gebruikt worden. Volgens par, 5.4. bezit $\underline{T}_{10}/(1 + \underline{T}_{10})$ een $\beta(n-1, n'-1)$-verdeling, waaruit de overschrijdingskansen:

$$k_\ell = P\left[\underline{T}_{10} \leq T_{10}\right] = P\left[\underline{u} \leq \frac{T_{10}}{1+T_{10}}\right] =$$
$$= \int_0^{T_{10}/(1+T_{10})} \frac{\Gamma(n+n'-2)}{\Gamma(n-1).\Gamma(n'-1)} \frac{u^{n-2}}{(1-u)^{n'-2}}\,du$$

en $$k_r = 1 - k_\ell$$

bepaald kunnen worden. Zijn $u(\varepsilon)$ en $u(1-\varepsilon)$ kritieke waarden uit deze β verdeling, dan is $T_{10}(\varepsilon) = \frac{u(\varepsilon)}{1-u(\varepsilon)}$ en $T_{10}(1-\varepsilon) = \frac{u(1-\varepsilon)}{1-u(1-\varepsilon)}$ en

$$\left(\frac{\sum_{i=1}^{n} (\underline{x}_i - \underline{x}_o)}{T_{10}(1-\varepsilon) \sum_{i=1}^{n'} (\underline{x}_i' - \underline{x}_o')} \; ; \; \frac{\sum_{i=1}^{n} (\underline{x}_i - \underline{x}_o)}{T_{10}(\varepsilon) \sum_{i=1}^{n'} (\underline{x}_i' - \underline{x}_o')} \right)$$

een betrouwbaarheidsinterval voor $\kappa = \frac{\beta}{\beta'}$ met onbetrouwbaarheid 2ε .

6.11. SUKHATME [25] geeft nog k-steekproeven toetsen voor exponentiële verdelingen $\Gamma(\alpha_i, \beta_i, 1)$ en de volgende hypothesen :

H_0 : $\alpha_1 = \ldots = \alpha_k = \alpha$ & $\beta_1 = \cdots = \beta_k = \beta$,

dus alle verdelingen dezelfde,

H_0' : $\beta_1 = \ldots = \beta_k = \beta$ en α_i willekeurig,

H_0'' : $\alpha_1 = \ldots = \alpha_k = \alpha$ onder de voorwaarde dat $\beta_1 = \cdots = \beta_k = \beta$.

6.12. In tabel 1 wordt nog een overzicht gegeven van de situaties wat betreft α , β en γ , waarvoor in dit hoofdstuk schattingen, toetsen of betrouwbaarheidsintervallen werden besproken. Hierbij kan nog opgemerkt worden dat in de gevallen waarin α bekend is de waarnemingen $\underline{x}_i$ ook wel uit Γ -verdelingen met verschillende α_i afkomstig mogen zijn; de schattingen en toetsingsgrootheden worden steeds met $x_i - \alpha_i$ berekend.

<u>Tabel 1.</u> Overzicht van de situaties waarvoor schattingen, toetsen of betrouwbaarheidsintervallen in verband met $\Gamma(\alpha, \beta, \gamma)$-verdelingen behandeld werden.

α	β	γ	<u>Paragraaf</u>
schatting	schatting	schatting	6.1
bekend	schatting	schatting	6.2
bekend	schatting voor β en voor $1/\beta$ toets voor $H_0: \beta = \beta_0$ betr. interval voor β	bekend, eventueel verschillende γ_i	6.3
schatting toets voor $H_0: \alpha = \alpha_0$ betr. interval	bekend	bekend, eventueel verschillende γ_i	6.4
onbekend	schatting en betr. interval toets voor $H_0: \beta = \beta_0$	1	6.5
meer steekproeven met bekende of onbekende verschillende α_i	schatting en betr. interval toets voor $H_0: \beta = \beta_0$	1	6.6
schatting en betr. interval toets voor $H_0: \alpha = \alpha_0$	onbekend	1	6.7
Bij levensduurproeven uit de kleinste r van n waarnemingen:			
bekend	schatting en betr. interval toets voor $H_0: \beta = \beta_0$	1	6.8
Bij vergelijking van twee of meer exponentiële verdelingen:			
toets voor $H_0: \alpha = \alpha' + \delta_0$ betr. interval voor $\delta = \alpha - \alpha'$	$\beta = \beta'$ of $\beta/n = \beta'/n'$ bekend	$\gamma = \gamma' = 1$	6.9
α en α' onbekend	toets voor $H_0: \beta = \kappa_0 \beta'$ betr. interval voor $\kappa = \frac{\beta}{\beta'}$	$\gamma = \gamma' = 1$	6.10
toets voor $H_0: \alpha_1 = \ldots = \alpha_k = \alpha$ &	$\beta_1 = \ldots = \beta_k = \beta$	$\gamma_1 = \ldots = \gamma_k = 1$	6.11
$= \alpha_1 = \ldots = \alpha_k$ onbekend	toets voor $H_0: \beta_1 = \ldots = \beta_k = \beta$ gegeven	$\gamma_1 = \ldots = \gamma_k = 1$	
toets voor $H_0: \alpha_1 = \ldots = \alpha_k = \alpha$	$\left.\beta_1 = \ldots = \beta_k = \beta\right\}$ onbekend	$\gamma_1 = \ldots = \gamma_k = 1$	

7. Appendix.

7.1. (par.3.2) Heeft $\underline{x}$ een $\Gamma(\alpha, \beta, \gamma)$ -verdeling, dan is de karakteristieke functie van $(\underline{x} - \alpha - \beta\gamma)/(\beta\sqrt{\gamma})$ gelijk aan (zie [3])

$$\varphi(t) = \mathcal{E}\, e^{it(\underline{x}-\alpha-\beta\gamma)/(\beta\sqrt{\gamma})} =$$

$$= \frac{1}{\beta^{\gamma}\Gamma(\gamma)} \int_{\alpha}^{\infty} e^{it(u-\alpha-\beta\gamma)/(\beta\sqrt{\gamma}) - (u-\alpha)/\beta} (u-\alpha)^{\gamma-1}\, du .$$

Stellen we hierin $v = \frac{u-\alpha}{\beta}$ en dus $du = \beta\, dv$, dan wordt

$$\varphi(t) = \frac{1}{\Gamma(\gamma)} \int_0^{\infty} e^{it(v/\sqrt{\gamma} - \sqrt{\gamma}) - v}\, v^{\gamma-1}\, dv =$$

$$= \frac{e^{-it\sqrt{\gamma}}}{\Gamma(\gamma)} \int_0^{\infty} e^{-v(1-it/\sqrt{\gamma})}\, v^{\gamma-1}\, dv = e^{-it\sqrt{\gamma}} / (1-it/\sqrt{\gamma})^{\gamma} =$$

$$= e^{-it\sqrt{\gamma} - \gamma \ln(1-it/\sqrt{\gamma})} = e^{-it\sqrt{\gamma} - \gamma[-it/\sqrt{\gamma} - \frac{1}{2}(it/\sqrt{\gamma})^2 - \frac{1}{3}(it/\sqrt{\gamma})^3 - \dots]} =$$

$$= e^{-\frac{1}{2}t^2 - \frac{1}{3}it^3/\sqrt{\gamma} + \dots} .$$

Hieruit volgt

$$\lim_{\gamma \to \infty} \varphi(t) = e^{-\frac{1}{2}t^2} ,$$

de karakteristieke functie van de N(0;1)-verdeling.

7.2. (par.5.1) $\underline{x}$ heeft een $\Gamma(\alpha, \beta, \gamma)$ -verdeling, dus:

$$P[\underline{x} \le x] = \frac{1}{\beta^{\gamma}\Gamma(\gamma)} \int_{\alpha}^{x} e^{-(u-\alpha)/\beta} (u-\alpha)^{\gamma-1}\, du .$$

Stel nu $x + a = y$ en $u + a = v$, dan geldt:

$$P[\underline{y} \le y] = P[\underline{x} \le y - a] = \frac{1}{\beta^{\gamma}\Gamma(\gamma)} \int_{\alpha}^{y-a} e^{-(u-\alpha)/\beta} (u-\alpha)^{\gamma-1}\, du =$$

$$= \frac{1}{\beta^{\gamma}\Gamma(\gamma)} \int_{\alpha+a}^{y} e^{-(v-\alpha-a)/\beta} (v-\alpha-a)^{\gamma-1}\, dv ,$$

waaruit volgt dat $\underline{y} = \underline{x} + a$ een $\Gamma(\alpha+a, \beta, \gamma)$ -verdeling heeft. Stel verder bx=z en bu=t $(b > 0)$, dan is

$$P[\underline{z} \leq z] = P\left[\underline{x} \leq \frac{z}{b}\right] = \frac{1}{\beta^{\gamma}\Gamma(\gamma)} \int_{\alpha}^{\frac{z}{b}} e^{-\frac{u-\alpha}{\beta}} (u-\alpha)^{\gamma-1} du =$$

$$= \frac{1}{\beta^{\gamma}\Gamma(\gamma)} \int_{b\alpha}^{z} e^{-\frac{t-b\alpha}{b\beta}} \left(\frac{t-b\alpha}{b}\right)^{\gamma-1} \frac{1}{b}\, dt =$$

$$= \frac{1}{(b\beta)^{\gamma}\Gamma(\gamma)} \int_{b\alpha}^{z} e^{-\frac{t-b\alpha}{b\beta}} (t-b\alpha)^{\gamma-1} dt,$$

dus $\underline{z} = b\underline{x}$ heeft een $\Gamma(b\alpha, b\beta, \gamma)$-verdeling.

7.3. (par.5.2). De karakteristieke functie van $\underline{x}_j$ met een $\Gamma(\alpha_j, \beta, \gamma_j)$-verdeling is

$$\varphi_{\underline{x}_j}(t) = \mathcal{E}\, e^{it\underline{x}_j} = \int_{\alpha_j}^{\infty} \frac{1}{\beta^{\gamma_j}\Gamma(\gamma_j)}\, e^{itx_j - \frac{x_j-\alpha_j}{\beta}} (x_j-\alpha_j)^{\gamma_j-1} dx_j .$$

Door hierin $(x_j - \alpha_j)/\beta = u$ te stellen wordt dit:

$$\varphi_{\underline{x}_j}(t) = \frac{1}{\Gamma(\gamma_j)} \int_0^{\infty} e^{it\beta u + \alpha_j - u}\, u^{\gamma_j-1} du = e^{it\alpha_j}(1-it\beta)^{-\gamma_j}.$$

De karakteristieke functie van $\sum_{j=1}^{n} \underline{x}_j$ waarin de $\underline{x}_j$ onderling onafhankelijk zijn is

$$\varphi_{\sum_{j=1}^{n}\underline{x}_j}(t) = \mathcal{E}\, e^{it\sum_{j=1}^{n}\underline{x}_j} = \prod_{j=1}^{n} \mathcal{E}\, e^{it\underline{x}_j} = e^{it\sum_{j=1}^{n}\alpha_j}(1-it\beta)^{-\sum_{j=1}^{n}\gamma_j},$$

dus $\sum_{j=1}^{n} \underline{x}_j$ bezit een $\Gamma\left(\sum_{j=1}^{n}\alpha_j, \beta, \sum_{j=1}^{n}\gamma_j\right)$-verdeling.

7.4. (par. 5.3. en 5.4.) $\underline{x}_1$ en $\underline{x}_2$ zijn onderling onafhankelijk $\Gamma(0, \beta, \gamma_1)$-resp. $\Gamma(0, \beta, \gamma_2)$ - verdeeld. De karakteristieke functie van $\underline{x}_1 + \underline{x}_2$ en $\underline{x}_1/(\underline{x}_1+\underline{x}_2)$ is

$$\varphi(t,s) = \mathcal{E}\, e^{it(\underline{x}_1+\underline{x}_2) + is\frac{\underline{x}_1}{\underline{x}_1+\underline{x}_2}} =$$

$$= \frac{1}{\beta^{\gamma_1+\gamma_2}\Gamma(\gamma_1)\Gamma(\gamma_2)} \int_0^{\infty}\int_0^{\infty} e^{it(x_1+x_2) + \frac{isx_1}{x_1+x_2} - \frac{x_1+x_2}{\beta}}\, x_1^{\gamma_1-1}\, x_2^{\gamma_2-1}\, dx_1\, dx_2 .$$

Stel nu $\begin{cases} x_1/(x_1+x_2) = y \\ x_1+x_2 = z \end{cases}$ of $\begin{cases} x_1=yz \\ x_2=z(1-y) \end{cases}$

dan is de determinant van JACOBI voor deze transformatie:

$$\begin{vmatrix} \frac{\partial x_1}{\partial y} & \frac{\partial x_1}{\partial z} \\ \frac{\partial x_2}{\partial y} & \frac{\partial x_2}{\partial z} \end{vmatrix} = \begin{vmatrix} z & y \\ -z & 1-y \end{vmatrix} = z,$$

terwijl 't integratie gebied voor y en z 't zelfde is als voor x_1 en x_2 (1e kwadraat van (y;z) vlak). De karakteristieke functie gaat daarom bij deze transformatie over in

$$\varphi(t,s) = \frac{1}{\beta^{\gamma_1+\gamma_2}\Gamma(\gamma_1)\Gamma(\gamma_2)} \int_0^\infty\int_0^\infty e^{itz+isy-\frac{z}{\beta}}(yz)^{\gamma_1-1} z^{\gamma_2-1}(1-y)^{\gamma_2-1} z\,dy\,dz =$$

$$= \int_0^\infty \frac{e^{itz-\frac{z}{\beta}} z^{\gamma_1+\gamma_2-1}\,dz}{\beta^{\gamma_1+\gamma_2}\Gamma(\gamma_1+\gamma_2)} \cdot \int_0^\infty \frac{\Gamma(\gamma_1+\gamma_2)\,e^{isy}}{\Gamma(\gamma_1).\Gamma(\gamma_2)} y^{\gamma_1-1}(1-y)^{\gamma_2-1}\,dy =$$

$$= \varphi(t).\varphi(s),$$

waarin $\varphi(t)$ de karakteristieke functie van $\underline{z}=\underline{x}_1+\underline{x}_2$ met een $\Gamma(0,\beta,\gamma_1+\gamma_2)$ -verdeling is en $\varphi(s)$ van $\underline{y} = \underline{x}_1/(\underline{x}_1+\underline{x}_2)$ met een $\beta(\gamma_1,\gamma_2)$ -verdeling. De grootheden $\underline{y}$ en $\underline{z}$ zijn dus onderling onafhankelijk.

7.5. (par.5.3.) De grootheden $\underline{x}_j$ (j=1, ..., n) zijn onderling onafhankelijk verdeeld volgens resp. een $\Gamma(0,\beta,\gamma_j)$ -verdeling. $F(x_1, \ldots, x_n)$ is een homogene functie van $x_1, \ldots, x_n$ van de graad 0, d.w.z.

$$F(kx_1,\ldots,kx_n) = F(x_1,\ldots,x_n).$$

De karakteristieke functie van $\sum_{j=1}^n \underline{x}_j$ en $\underline{F} = F(\underline{x}_1, \ldots, \underline{x}_n)$ is:

$$\varphi(t,s) = \int_0^\infty \frac{e^{it\sum_{j=1}^n x_j + isF - \frac{\sum_{j=1}^n x_j}{\beta}}}{\beta^{\sum_{j=1}^n \gamma_j}\prod_{j=1}^n \Gamma(\gamma_j)} \prod_{j=1}^n x_j^{\gamma_j-1}\,dx_1\ldots dx_n.$$

De transformatie

$$\begin{cases} y_1 = x_1/\sum_{j=1}^n x_j \\ \vdots \\ y_{n-1} = x_{n-1}/\sum_{j=1}^n x_j \\ y_n = \sum_{j=1}^n x_j \end{cases} \quad \text{dus} \quad \begin{cases} x_1 = y_n\,y_1 \\ \vdots \\ x_{n-1} = y_n\,y_{n-1} \\ x_n = y_n(1-\sum_{j=1}^{n-1} y_j) \end{cases}$$

met de determinant van JACOBI = y_n^{n-1}, waarbij het integratiegebied voor $y_1, \ldots, y_n$ hetzelfde is als voor $x_1, \ldots, x_n$, voert deze functie over in

$$\varphi(t,s) = \int_0^\infty \ldots \int_0^\infty \frac{e^{ity_n + isF(y_1,\ldots,y_{n-1},1) - \frac{y_n}{\beta}}}{\beta^{\sum_{j=1}^n \gamma_j} \prod_{j=1}^n \Gamma(\gamma_j)} \prod_{j=1}^{n-1} (y_n y_j)^{\gamma_j - 1} \left\{ y_n \left(1 - \sum_{j=1}^{n-1} y_j\right) \right\}^{\gamma_n - 1} y_n^{n-1}\, dy_1 \ldots dy_n =$$

$$= \int_0^\infty \frac{e^{ity_n - \frac{y_n}{\beta}}\, y_n^{\sum_{j=1}^n \gamma_j - 1}}{\beta^{\sum_{j=1}^n \gamma_j} \Gamma\left(\sum_{j=1}^n \gamma_j\right)} dy_n \cdot \int_0^\infty \ldots \int_0^\infty \frac{\Gamma\left(\sum_{j=1}^n \gamma_j\right)}{\prod_{j=1}^n \Gamma(\gamma_j)} e^{isF(y_1,\ldots,y_{n-1},1)} \prod_{j=1}^{n-1} y_j^{\gamma_j - 1} \left(1 - \sum_{j=1}^{n-1} y_j\right)^{\gamma_n - 1} dy_1 \ldots dy_{n-1} =$$

$$= \varphi(t) . \varphi(s),$$

waarin $\varphi(t)$ de karakteristieke functie van $\underline{y}_n = \sum_{j=1}^n \underline{x}_j$ is en $\varphi(s)$ die van $F(\underline{y}_1, \ldots, \underline{y}_{n-1}, 1) = F(\underline{x}_1, \ldots, \underline{x}_n)$; dus $\underline{F}$ en $\sum_{j=1}^n \underline{x}_j$ zijn onafhankelijk verdeeld.

7.6. (par. 5.6.) (Zie [21]). De waarnemingen $\underline{x}_j$ ($j=1, \ldots, n$) zijn onderling onafhankelijk $\Gamma(\alpha, \beta, 1)$- verdeeld.

$$\underline{x}_0 \overset{def}{=} \min(\underline{x}_1, \ldots, \underline{x}_n) \quad \text{en} \quad \underline{y} \overset{def}{=} \sum_{j=1}^n (\underline{x}_j - \underline{x}_0).$$

De karakteristieke functie van $\underline{x}_0$ en $\underline{y}$,

$$\varphi(t,s) = \mathcal{E}\, e^{it\underline{y} + is\underline{x}_0},$$

is nu een integraal over een gebied in de $(x_1, \ldots, x_n)$-ruimte, dat uit n niet-overlappende gedeelten bestaat; in elk deel is nl. één van de x_j (die dan de rol van x_0 speelt) kleiner dan alle andere. Daar alle $\underline{x}_i$ dezelfde verdeling bezitten en op dezelfde manier tot $\underline{x}_0$ en $\underline{y}$ bijdragen is de integraal over elk van deze deelgebieden dezelfde. Het is dus voldoende de integraal over het deel waar x_1 kleiner is dan $x_2, \ldots, x_n$ te berekenen en met n te vermenigvuldigen:

$$\varphi(t,s) = n \int \ldots \int_{x_1 \leq \min(x_2,\ldots,x_n)} \frac{1}{\beta^n} e^{it\sum_{j=2}^n (x_j - x_1) + isx_1 - \sum_{j=1}^n (x_j - \alpha)/\beta}\, dx_1 \ldots dx_n .$$

Stel:

$$\begin{cases} x_1 - \alpha = z_1 \\ x_j - x_1 = z_j, \quad j = 2, \ldots, n \end{cases} \quad \text{of} \quad \begin{cases} x_1 = z_1 + \alpha \\ x_j = z_1 + z_j + \alpha . \end{cases}$$

Bij deze transformatie is de determinant van JACOBI = 1, terwijl het integratiegebied in de $(x_1, \ldots, x_n)$-ruimte overgaat in $(0 \leq z_1 \leq \infty ; \ldots ; 0 \leq z_n \leq \infty)$ in de $(z_1, \ldots, z_n)$-ruimte, dus:

$$\varphi(t,s)=\frac{n}{\beta^n}\int_0^\infty\cdots\int_0^\infty e^{it\sum_{j=2}^n z_j-\sum_{j=2}^n z_j/\beta}\,e^{is(z_1+\alpha)-\frac{nz_1}{\beta}}dz_1\;dz_2\ldots dz_n=$$
$$=\varphi(t).\varphi(s),$$

waarin $\varphi(s)$ de karakteristieke functie van $\underline{z}_1+\alpha=\underline{x}_0$ met een $\Gamma(\alpha,\frac{\beta}{n},1)$ -verdeling is en $\varphi(t)$ van $\sum_{j=2}^n \underline{z}_j=\sum_{j=1}^n(\underline{x}_j-\underline{x}_0)$ met een $\Gamma(0,\beta,n-1)$-verdeling. $\underline{x}_0$ en $\sum_{j=1}^n(\underline{x}_j-\underline{x}_0)$ zijn dus onafhankelijk.

7.7. (par.5.7.) Is $T(x_1,\ldots,x_n)$ onafhankelijk van de keuze van de oorsprong, dan kan T als functie van x_i-x_0 $(i=1,\ldots,n)$ geschreven worden. Volkomen analoog aan het betoog in par.7.6. is dan ook nu te bewijzen dat de simultane karakteristieke functie van $\underline{T}$ en $\underline{x}_0$ het product is van de karakteristieke functies van $\underline{T}$ en $\underline{x}_0$ afzonderlijk, dus dat $\underline{T}$ en $\underline{x}_0$ onafhankelijk verdeeld zijn.

7.8. (par.5.8.). Is $G(x_1,\ldots,x_n)$ onafhankelijk van de keuze van de oorsprong, als T in par. 5.7., en homogeen van de graad 0 als F in par. 5.3., dan kan G als functie van x_j-x_0 $(j=1,\ldots,n)$ geschreven worden en geldt:

$$G(k(x_1-x_0),\ldots,k(x_n-x_0))=G(x_1-x_0,\ldots,x_n-x_0).$$

De simultane onafhankelijkheid van $\underline{G}$, $\underline{x}_0$ en $\sum_{i=1}^n(\underline{x}_i-\underline{x}_0)$ is dan te bewijzen door op de karakteristieke functie

$$\varphi(s,t,u)=\mathcal{E}\,e^{is\underline{G}+it\underline{x}_0+iu\sum_{i=1}^n(\underline{x}_i-\underline{x}_0)}$$

achtereenvolgens transformaties als in par. 7.6. en par.7.5. toe te passen. $\varphi(s,t,u)$ blijkt dan het product van de karakteristieke functies van $\underline{G}$, $\underline{x}_0$ en $\sum_{i=1}^n(\underline{x}_i-\underline{x}_0)$ te zijn.

7.9. (par.6.2.). De aannemelijkheidsfunctie voor een steekproef $\underline{x}_1,\ldots,\underline{x}_n$ uit een $\Gamma(0,\beta,\gamma)$ -verdeling is:

$$L(\beta,\gamma)=\prod_{i=1}^n g(x_i;0,\beta,\gamma)=\frac{e^{-\frac{\sum_{i=1}^n x_i}{\beta}}\prod_{i=1}^n x_i^{\gamma-1}}{\beta^{n\gamma}(\Gamma(\gamma))^n}$$

dus: $$\ln L(\beta,\gamma)=-\frac{\sum_{i=1}^n x_i}{\beta}-n\gamma\ln\beta-n\ln\Gamma(\gamma)+(\gamma-1)\sum_{i=1}^n\ln x_i,$$

$$\frac{\partial\ln L}{\partial\beta}=\frac{\sum_{i=1}^n x_i}{\beta^2}-\frac{n\gamma}{\beta}$$

en:

$$\frac{\partial \ln L}{\partial \gamma} = -n \ln \beta - n \frac{d}{d\gamma} \ln \Gamma(\gamma) + \sum_{i=1}^{n} \ln x_i .$$

De meest aannemelijke schattingen $\hat{b}$ en $\hat{c}$ voor β en γ maken deze laatste twee uitdrukkingen nul, dus:

$$\hat{b} = \frac{\bar{x}}{\hat{c}}$$

en

$$\ln \hat{c} - \frac{\Gamma(\hat{c})}{\Gamma'(\hat{c})} = \ln \bar{x} - \frac{1}{n} \sum_{i=1}^{n} \ln x_i .$$

Uit deze laatste vergelijking kan $\hat{c}$ alleen opgelost worden met behulp van tabellen voor de functie $\ln t - \Gamma(t)/\Gamma'(t)$. Een benadering ontstaat door van de reeksontwikkeling voor $\Gamma(t)/\Gamma'(t)$

$$\Gamma(t)/\Gamma'(t) = \ln t - \frac{1}{2t} + \cdots$$

alleen de eerste twee termen te gebruiken, dit geeft:

$$\frac{1}{2\hat{c}} = \ln \bar{x} - \frac{1}{n} \sum_{i=1}^{n} \ln x_i \quad \text{of} \quad \hat{c} = \frac{1}{2\left(\ln \bar{x} - \frac{1}{n} \sum_{i=1}^{n} \ln x_i\right)} .$$

7.10. (par.6.2.) Het steekproef experiment omvatte in totaal 10 steekproeven van elk 20 waarnemingen en 5 steekproeven van elk 40 waarnemingen bij verschillende waarden van β en γ ($\alpha = 0$). De waarnemingen werden gevonden door aselecte getallen tussen 0 en 1 via de tabellen van SALVOSA of K. PEARSON van de verdelingsfunctie $G(x; o, \beta, \gamma)$, met de gewenste waarden voor β en γ (dus α_3 resp. p) te transformeren tot een waarneming uit de $\Gamma(o, \beta, \gamma)$ -verdeling. Uit elke steekproef werden op 3 manieren schattingen voor β en γ berekend; nl.:

1) meest aannemelijke schattingen met behulp van de tabellen van GREENWOOD en DURAND ($\hat{b}$ en $\hat{c}$);

2) benaderingen hiervoor volgens de formules (6.2;2)($b_1$enc_1);

3) schattingen uit de momenten volgens de formules (6.2;1) (b_2, c_2).

De resultaten zijn verzameld in tabel 2. Ze laten zien dat de afwijking tussen de werkelijke waarde van de parameter en de schattingen hiervoor bij alle drie methodes aanzienlijk kan zijn.

Tabel 2. Schattingen voor β en γ uit steekproefexperimenten.

steekproef-grootte	$\Gamma(0,\beta,\gamma)$		meest aannem. schattingen		benad.schatt.		mom.meth.	
	β	γ	$\hat{b}$	$\hat{c}$	$\hat{b}_1$	$\hat{c}_1$	b_2	c_2
20	4	25	2,45	39,85	2,47	39,57	2,35	41,58
20	4	25	3,96	26,34	3,98	26,20	4,26	24,50
20	4	4	2,83	5,50	2,91	5,34	2,95	5,27
20	4	4	5,88	2,81	6,24	2,65	6,42	2,58
20	1	100	0,53	192,47	0,54	190,57	0,55	185,58
20	1	100	1,28	80,81	1,28	80,42	1,28	80,70
20	1	16	1,20	13,75	1,22	13,58	1,15	14,37
20	1	16	1,27	12,07	1,29	11,91	1,37	11,19
20	1	6,25	0,93	7,54	0,95	7,37	0,93	7,56
20	1	6,25	0,74	7,78	0,75	7,62	0,87	6,58
40	1	0,2	0,39	0,52	0,49	0,41	0,56	0,36
40	1	0,4	0,69	0,58	0,85	0,47	0,59	0,68
40	1	0,6	0,97	0,73	1,16	0,61	1,07	0,66
40	1	0,8	0,91	1,08	1,04	0,94	0,80	1,23
40	1	1,0	0,75	1,40	0,83	1,26	0,74	1,42

7.11.(par.6.3.) Hebben $\underline{x}_1, \ldots, \underline{x}_n$ een $\Gamma(0,\beta,\gamma)$ -verdeling dan bezit $\underline{\bar{x}}$ een $\Gamma(0, \frac{\beta}{n}, n\gamma)$ -verdeling, dus

$$\mathcal{E}\left(\frac{1}{\underline{b}'}\right) = \mathcal{E}\,\frac{n\gamma-1}{n\underline{\bar{x}}} = \int_0^\infty \frac{n\gamma-1}{\left(\frac{\beta}{n}\right)^{n\gamma}\Gamma(n\gamma)} \cdot \frac{1}{n\bar{x}} \cdot e^{-\frac{n\bar{x}}{\beta}}\, \bar{x}^{\,n\gamma-1}\, d\bar{x} =$$

$$= \frac{1}{\beta} \cdot \int_0^\infty \frac{1}{\left(\frac{\beta}{n}\right)^{n\gamma-1}\Gamma(n\gamma-1)}\, e^{-\frac{n\bar{x}}{\beta}}\, \bar{x}^{\,n\gamma-2}\, d\bar{x} = \frac{1}{\beta}\,.$$

7.12. (par.6.7.) De waarden α_0 voor α die op grond van het gevonden waarnemingsmateriaal door de besproken toets niet verworpen worden, vormen een betrouwbaarheidsinterval voor α met de onbetrouwbaarheid van de toets. Dit zijn dus de waarden α_0 waarvoor:

$$T_7(\varepsilon) \leq n(x_0 - \alpha_0) / \sum_{i=1}^{n} (x_i - \alpha_0)$$

of:

$$T_7(\varepsilon) \sum_{i=1}^{n} (x_i - \alpha_0) \leq n(x_0 - \alpha_0)$$

$$T_7(\varepsilon) \cdot \bar{x} - x_0 \leq \alpha_0 (T_7(\varepsilon) - 1)$$

$$\frac{x_0 - \bar{x}\, T_7(\varepsilon)}{1 - T_7(\varepsilon)} \leq \alpha_0 .$$

α is in ieder geval kleiner dan de kleinste waarneming:

$$P[\alpha \leq \underline{x}_0] = 1 ,$$

dus:

$$\left(\frac{x_0 - \bar{x}\, T_7(\varepsilon)}{1 - T_7(\varepsilon)} \; ; \; \underline{x}_0 \right)$$

is een betrouwbaarheidsinterval voor α met onbetrouwbaarheid ε .

Litteratuur

In deze litteratuurlijst zijn ook artikelen opgenomen, welke in dit rapport niet ter sprake zijn gekomen.

[1] H. Akaike (1956) On the distribution of the product of two Γ-distributed variables. Ann. Inst.Math.Tokyo, 8, pp. 53-54.

[2] G.D. Berndt (1958) Power functions of the Γ-distribution. A.M.S., 29, pp. 302-306.

[3] H. Cramer (1946) Mathematical Methods of Statistics. Princeton, University Press.

[3a] D.G. Chapman (1956) Estimating the parameters of a truncated gamma distribution. Annals 27, pp. 498-506.

[4] D. van Dantzig (1947) Kadercursus Mathematische Statistiek, hoofdstuk 3. Amsterdam, Math.Centrum.

[5] B. Epstein, M. Sobel (1953) Life testing. J.A.S.A., 48, pp.486-502.

[6] B. Epstein, M. Sobel (1954) Some theorems relevant to life testing from an exponential distribution. A.M.S., 25, pp.373-381.

[7] L.A. Goodman (1951-52) On the Poisson-Gamma distribution problem. Ann.Inst.Stat.Math.Tokyo, 3, pp.123-125.

[8] J.A. Greenwood, D. Durand (1958) Aids for fitting the Pearson type III curve by maximum likelihood. Nog niet gepubliceerd, abstract in A.M.S., 29, p. 1280.

[9] M.V. Jambunathan (1954) Some properties of β - and γ - distributions. A.M.S., 25, pp.401-405.

[10] M.G. Kendall (1947) The advanced theory of statistics I & II. Ch. Griffin & Co. Lmd., London.

[11] A.S. Krishnamoorthy & M.Parthasarathy (1951) A multivariate Gamma-type distribution A.M.S., 22, pp.549-557.

[12] S. Kullback (1936) The distribution laws of the difference and quotients of variables independently distributed in Pearson type III laws. A.M.S., 7, pp.51-53.

[13] R.G. Laha (1954) On a characterisation of the gamma distribution. A.M.S., 25, pp.784-787.

[14] R.G. Laha (1956) On some properties of the normal and the gamma distribution. Proc.Ann. Math.Soc., 7, pp.172-174.

[15] E. Lukacs (1955) A characterisation of the gamma distribution. A.M.S., 26, pp.319-324.

[16] E. Lukacs (1956) Characterisation of populations by properties of suitable statistics. Third Berkeley Symposium on Math. Stat. and Prob., II, pp.195-215.

[17] Seiji Nabeya (1950) On a relation between exponential law and Poisson's law. Ann.Inst.Stat. Math. Tokyo, 2, pp. 13-16.

[18] A.C. Olshen (1938) Transformations of the Pearson type III distribution. A.M.S., 9, pp. 176-200.

[19] K. Pearson (1946) Tables of the incomplete Γ-function. Cambridge, University Press.

[20] E.S. Pearson, H.O.Hartley(1954) Biometrika tables for Statisticians Vol. I. Cambridge University Press, p. 122-131.

[21] E.J.G. Pitman (1937) The "closest" estimates of statistical parameters. Proc.Cambr.Phil.Soc., 33, pp. 212-222.

[22] C.R. Rao (1952) Advanced Statistical Methods in Biometric Research, Chapman & Hall, London.

[23] L.R. Salvosa (1930) Tables of Pearson's type III function. A.M.S., 1, pp. 191-198.

[24] K.C. Seal (1957) On a characterisation of gamma distributions. Calcutta Stat.Assoc. Bull., 7, pp. 60-72.

[25] P.V. Sukhatme (1936) Analysis of k samples from exponential populations. Stat.Research Memoirs, 1, pp. 94-112.

[26] E.T. Whittaker, G.N.Watson(1946) A course of modern analysis. Cambridge, University Press.