FORMULARIUM: ELEMENTAIRE STATISTIEK

HULPFORMULES

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \text{ als } |x| < 1.$$
$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$$
$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$
$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$
$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$$
$$\left(1+\frac{a}{n}\right)^n \stackrel{n\to\infty}{\longrightarrow} e^a$$
$$\Gamma(t) \sim \left(\frac{t-1}{e}\right)^{t-1} \sqrt{2\pi(t-1)}$$

DISCRETE KANSMODELLEN

naam	$f(k) = \mathrm{P}(X = k)$	Ω	$E[X]$	$\mathrm{Var}[X]$	$M_X(t)$
discreet uniform	$\frac{1}{n}$	$k = 1, 2, \ldots, n$	$\frac{n+1}{2}$	$\frac{(n+1)(n-1)}{12}$	$\frac{e^t(e^{tn}-1)}{n(e^t-1)}$
Bernoulli	$p^k(1-p)^{1-k}$	$k = 0, 1$	p	$p(1-p) = pq$	$q + pe^t$
binomiaal	$\binom{n}{k} p^k(1-p)^{n-k}$	$k = 0, \ldots, n$	np	$np(1-p) = npq$	$(q + pe^t)^n$
Poisson	$e^{-\alpha}\frac{\alpha^k}{k!}$	$k = 0, 1, 2, \ldots$	α	α	$e^{\alpha(e^t-1)}$
geometrisch	$(1-p)^k p$	$k = 0, 1, 2 \ldots$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$	$\begin{cases} \frac{p}{1-(1-p)e^t} & t < -\ln(1-p) \\ \infty & t \geq -\ln(1-p) \end{cases}$
negatief binomiaal	$\binom{k+r-1}{k} p^r(1-p)^k$	$k = 0, 1, 2 \ldots$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$	$\begin{cases} \left(\frac{p}{1-(1-p)e^t}\right)^r & t < -\ln(1-p) \\ \infty & t \geq -\ln(1-p) \end{cases}$
hypergeometrisch	$\frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$	$k = 0, \ldots, n$	$\frac{rn}{N}$	$\frac{rn(N-r)(N-n)}{N^2(N-1)}$	/

Opmerkingen:
Soms gebruikt men ook de notaties $\mathrm{Geom}(\theta)$ en $\mathrm{NB}(r, \theta)$ waarbij $\theta = 1 - p$ de kans op mislukking.

CONTINUE KANSMODELLEN

naam	dichtheid $f(x)$	Ω	$E[X]$	Var[X]	$M_X(t)$
continu uniform	$\frac{1}{b-a}$	$x \in [a,b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{1}{b-a}\frac{1}{t}\left(e^{tb}-e^{ta}\right)$
$\mathcal{N}(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$, $\mu \in \mathbb{R}, \sigma > 0$	$x \in \mathbb{R}$	μ	σ^2	$e^{\mu t + \frac{\sigma^2 t^2}{2}}$
Expo(α)	$\alpha e^{-\alpha x}$, $\alpha > 0$	$x > 0$	$\frac{1}{\alpha}$	$\frac{1}{\alpha^2}$	$\frac{\alpha}{\alpha-t}$, $t < \alpha$
χ^2_n	$\frac{2^{-n/2}}{\Gamma(\frac{n}{2})}x^{\frac{n}{2}-1}e^{-x/2}$, $n > 0$	$x > 0$	n	$2n$	$\begin{cases}\frac{1}{(1-2t)^{\frac{n}{2}}} & t < \frac{1}{2}\\ \infty & t \geqslant \frac{1}{2}\end{cases}$
Gamma(γ,β)	$\frac{x^{\gamma-1}e^{\frac{-x}{\beta}}}{\beta^\gamma\Gamma(\gamma)}$, $\gamma,\beta > 0$	$x > 0$	$\gamma\beta$	$\gamma\beta^2$	$\begin{cases}(1-\beta t)^{-\gamma} & t < \frac{1}{\beta}\\ \infty & t \geq \frac{1}{\beta}\end{cases}$
t_n	$\frac{\Gamma((n+1)/2)}{\Gamma(n/2)\sqrt{\pi n}}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2}$, $n > 0$	$x \in \mathbb{R}$	0 $(n > 1)$	$\frac{n}{n-2}$ $(n > 2)$	/
Cauchy	$\frac{1}{\pi}\frac{1}{1+x^2}$	$x \in \mathbb{R}$	/	/	/
$F_{m,n}$	$\frac{m^{\frac{m}{2}}n^{\frac{n}{2}}\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\frac{x^{\frac{m}{2}-1}}{(n+mx)^{\frac{n+m}{2}}}$, $m,n > 0$	$x > 0$	$\frac{n}{n-2}$	$\frac{2n^2(m+n-2)}{m(n-4)(n-2)^2}$	/
lognormale	$\frac{1}{\sqrt{2\pi}\sigma x}e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2}$, $\mu \in \mathbb{R}, \sigma > 0$	$x \in \mathbb{R}$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)$	/
bivariaat normaal	$f(x,y) = \frac{1}{2\pi\sqrt{det(\Sigma)}}e^{-\frac{1}{2}z^t\Sigma^{-1}z}$ $\mu^t = (\mu_X\ \mu_Y)$, $z^t = (x\ y) - \mu^t$ en $\Sigma = \begin{pmatrix}\sigma_X^2 & \rho\sigma_X\sigma_Y\\ \rho\sigma_X\sigma_Y & \sigma_Y^2\end{pmatrix}$	$x,y \in \mathbb{R}$	$E[X] = \mu_X$, $E[Y] = \mu_Y$	$\mathrm{Var}[X] = \sigma_X^2$, $\mathrm{Var}[Y] = \sigma_Y^2$ $Cov(X,Y) = \rho\sigma_X\sigma_Y$	$M_X(s) = e^{s^t\mu+\frac{1}{2}s^t\Sigma s}$

Opmerkingen: $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}dx, \quad t > 0$

$\mathcal{B}(s,t) = \frac{\Gamma(s)\Gamma(t)}{\Gamma(s+t)}, \quad s,t > 0$

Als $X \sim \mathcal{N}\left(\mu_X,\sigma_X^2\right)$ en $Y \sim \mathcal{N}\left(\mu_Y,\sigma_Y^2\right)$ dan $X \mid Y = y \sim \mathcal{N}\left(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y), \sigma_X^2(1-\rho^2)\right)$

HYPOTHESETOETSEN

Tabel 1: normaliteitstest

n	$\alpha = 0.01$	$\alpha = 0.05$
10	.880	.918
15	.911	.938
20	.929	.950
25	.941	.958
30	.949	.964
40	.960	.972
50	.966	.976
60	.971	.980
75	.976	.984
100	.981	.986
150	.987	.991
200	.990	.993

- $r = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1}+\frac{(s_2^2/n_2)^2}{n_2-1}}$ (Fisher-Behrens)
- **Kolmogorov-Smirnov** geeft verdeling van D_n onder H_0

$$\lim_{n\to\infty} P(\sqrt{n}D_n \leq d) = 1 - 2\sum_{k=1}^{\infty}(-1)^{k-1}\exp(-2k^2d^2) := H(d),\ \forall d > 0.$$

Voor $n \gg$: verwerp H_0 als $\sqrt{n}D_n \geq H^{-1}(1-\alpha)$

d	$H(d)$	d	$H(d)$	d	$H(d)$	d	$H(d)$
0.30	0.0000	0.75	0.3728	1.20	0.8878	1.80	0.9969
0.35	0.0003	0.80	0.4559	1.25	0.9121	1.90	0.9985
0.40	0.0028	0.85	0.5347	1.30	0.9319	2.00	0.9993
0.45	0.0126	0.90	0.6073	1.35	0.9478	2.10	0.9997
0.50	0.0361	0.95	0.6725	1.40	0.9603	2.20	0.9999
0.55	0.0772	1.00	0.7300	1.45	0.9702	2.30	0.9999
0.60	0.1357	1.05	0.7798	1.50	0.9778	2.40	1.0000
0.65	0.2080	1.10	0.8223	1.60	0.9880	2.50	1.0000
0.70	0.2888	1.15	0.8580	1.70	0.9938		

- $\frac{R_{XY}}{\sqrt{\frac{(1-R_{XY}^2)}{n-2}}} \sim t_{n-2} \qquad |H_0$
- Als α en β de werkelijke parameters zijn dan geldt

$$\frac{(\widehat{\beta}-\beta)\sqrt{\sum_{i=1}^{n}(x_i-\overline{x}_n)^2}}{s_\epsilon} \sim t_{n-2}$$

$$\frac{(\widehat{\alpha}-\alpha)\sqrt{\sum_{i=1}^{n}(x_i-\overline{x}_n)^2}}{s_\epsilon\sqrt{\overline{x}_n^2+\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x}_n)^2}} \sim t_{n-2}$$