FORMULARIUM: KANSREKENEN EN STATISTIEK I

HULPFORMULES

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

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	$(1-p)^k p$	$k = 0, 1, 2 \ldots$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$	$\frac{p}{1-(1-p)e^t}$
negatief binomiaal	$\binom{k+r-1}{k} p^r(1-p)^k$	$k = 0, 1, 2 \ldots$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{p}{1-(1-p)e^t}\right)^r$
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	$\frac{1}{b-a}$	$x \in [a,b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{1}{b-a}\frac{1}{t}\left(e^{tb}-e^{ta}\right)$
$\mathcal{N}(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$, $\mu \in \mathbb{R}, \sigma > 0$	$x \in \mathbb{R}$	μ	σ^2	$e^{\mu t + \frac{\sigma^2 t^2}{2}}$
$\mathcal{E}(\alpha)$	$\alpha e^{-\alpha x}$, $\alpha > 0$	$x > 0$	$\frac{1}{\alpha}$	$\frac{1}{\alpha^2}$	$\frac{\alpha}{\alpha - t}$, $t < \alpha$
χ^2_n	$\frac{2^{-n/2}}{\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}e^{-x/2}$, $n > 0$	$x > 0$	n	$2n$	$\begin{cases} \frac{1}{(1-2t)^{\frac{n}{2}}} & t < \frac{1}{2} \\ \infty & t \geqslant \frac{1}{2} \end{cases}$
$\Gamma_{\gamma,\beta}$	$\frac{x^{\gamma-1}e^{\frac{-x}{\beta}}}{\beta^\gamma \Gamma(\gamma)}$, $\gamma, \beta > 0$	$x > 0$	$\gamma\beta$	$\gamma\beta^2$	$\begin{cases} (1-\beta t)^{-\gamma} & t < \frac{1}{\beta} \\ \infty & t \geq \frac{1}{\beta} \end{cases}$
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 $\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}$	$x, y \in \mathbb{R}$	$E[X] = \mu_X$, $E[Y] = \mu_Y$	$\text{Var}[X] = \sigma_X^2$, $\text{Var}[Y] = \sigma_Y^2$ $Cov(X,Y) = \rho\sigma_X\sigma_Y$	$M_X(s) = e^{s^t\mu + \frac{1}{2}s^t\Sigma s}$

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

SCHATTEN VAN PARAMETERS

parameter	schatter	verdeling	BI
μ bij $\mathcal{N}(\mu,\sigma^2)$	$\bar{X}$	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$ $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$	$\left[\bar{x}-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \bar{x}+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$ $\left[\bar{x}-t_{n-1,1-\frac{\alpha}{2}}\dfrac{s}{\sqrt{n}}; \bar{x}+t_{n-1,1-\frac{\alpha}{2}}\dfrac{s}{\sqrt{n}}\right]$
σ^2 bij $\mathcal{N}(\mu,\sigma^2)$	S^2	$(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$	$\left[\frac{(n-1)s^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}, \frac{(n-1)s^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$
proportie p	$\hat{P}$	$\dfrac{\hat{P}-p}{\sqrt{\frac{p(1-p)}{n}}} \approx \mathcal{N}(0,1)$ $n\hat{P} \sim \mathcal{B}(n,p)$	$\left[\hat{p}-z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p}+z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$

Opmerkingen: $z_\alpha = \Phi^{-1}(\alpha)$
$t_{n,\alpha} = Q_T(\alpha)$ met $T \sim t_n$
$\chi^2_{n,\alpha} = Q_X(\alpha)$ met $X \sim \chi^2_n$

HYPOTHESETOETSEN

Hypothese	Mogelijke teststatistieken
$\mu = \mu_0$	$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$ $T = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$
$\sigma^2 = \sigma_0^2$	$X = \dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$
$p = p_0$	$Z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \approx \mathcal{N}(0,1)$ $n\hat{P} \sim \mathcal{B}(n, p_0)$
$\mu_X = \mu_Y$	$T = \frac{\bar{X}-\bar{Y}}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$ met $S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ $T = \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \sim t_r$ met $r = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1}+\frac{(s_2^2/n_2)^2}{n_2-1}}$ Definieer $D = X - Y$ en test $\mu_D = \mu_0$
$\sigma_X^2 = \sigma_Y^2$	$F = S_X^2/S_Y^2 \sim F_{n_1-1,n_2-1}$ opm: p-waarde= $2\mathrm{P}(F_{n_1-1,n_2-1} > f)$ als $f > F_{n_1-1,n_2-1,0.5}$ $2\mathrm{P}(F_{n_1-1,n_2-1} < f)$ als $f \leqslant F_{n_1-1,n_2-1,0.5}$
$p_1 = p_2$	$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}} \approx \mathcal{N}(0,1)$ Benaderend: $Z' = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1}+\frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} \approx \mathcal{N}(0,1)$
$\rho = 0$	$\frac{R\sqrt{n-2}}{\sqrt{1-R^2}} \sim t_{n-2}$
onafhankelijkheid	$\chi^2 \approx \chi^2_m$ met m=(rijen-1)(kolommen-1) χ^2=som over alle cellen van $\dfrac{(\text{geobserveerde waarde - verwachte waarde})^2}{\text{verwachte waarde}}$ waarin: verwachte waarde = $\dfrac{\text{rijsom}\times\text{kolomsom}}{n}$
normaliteit	correlatiecoëfficiënt van normale kwantielplot, zie Tabel 1

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