Formuleblad Statistiek deel 1 (2024-2025)

Gegeven is een steekproef met n uitkomsten $x_1, x_2, \ldots, x_n$.

Steekproefgemiddelde:

$$\overline{x} = \frac{\sum\limits_i x_i}{n} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Steekproefvariantie:

$$s^2 = \frac{\sum\limits_i (x_i - \overline{x})^2}{n} = \frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \ldots + (x_n - \overline{x})^2}{n} \qquad \text{(optie 1)}$$

$$s^2 = \frac{\left(\sum\limits_i x_i^2\right) - n \cdot \overline{x}^2}{n} = \frac{(x_1^2 + x_2^2 + \ldots + x_n^2) - n \cdot \overline{x}^2}{n} \qquad \text{(optie 2)}$$

Rekenregels kansrekening:

$$P(A \text{ of } B) = P(A) + P(B) - P(A \text{ en } B) \qquad \text{(optelregel)}$$

$$P(B) = 1 - P(\text{niet } B) \qquad \text{(complementregel)}$$

$$P(A \mid B) = \frac{P(A \text{ en } B)}{P(B)} \qquad \text{(conditionele kansen)}$$

Discrete en continue kansverdelingen:

	Discrete kansvariabelen	**Continue kansvariabelen**
Uitkomstenruimte:	Eindig / aftelbaar oneindig	Overaftelbaar oneindig
Toepassingen:	Tellen / categoriseren	Meten
Kansbegrip:	Kansfunctie $p(k) = P(X = k)$	Kansdichtheidsfunctie $f(x)$
CDF:	$F(k) = P(X \leq k) = \sum_{\ell : \ell \leq k} p(\ell)$	$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)\, dy$
Verwachtingswaarde:	$E[X] = \sum_k k \cdot P(X = k)$	$E[X] = \int x \cdot f(x)\, dx$
Variantie:	$\mathrm{Var}(X) = \sum_k (k - E[X])^2 \cdot P(X = k)$	$\mathrm{Var}(X) = \int (x - E[X])^2 \cdot f(x)\, dx$
Standaardafwijking:	$\sigma(X) = \sqrt{\mathrm{Var}(X)}$	$\sigma(X) = \sqrt{\mathrm{Var}(X)}$

Verwachtingswaarde en variantie van veelgebruikte kansverdelingen:

Verdeling	Kans(dichtheids)functie	CDF	$E(X)$	Var(X)
Discreet				
Uniform(a,b)	$p(k)=\frac{1}{b-a+1}$ $(k=a,a+1,\ldots,b)$	$F(k)=\begin{cases}0 & x<a\\ \frac{k-a+1}{b-a+1} & a\le k<b\\ 1 & k\ge b\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$
Binomiaal(n,p)	$p(k)=\binom{n}{k}p^k(1-p)^{n-k}$	$F(k)=\sum_{i=0}^{k}\binom{n}{i}p^i(1-p)^{n-i}$	np	$np(1-p)$
Poisson(λ)	$p(k)=e^{-\lambda}\cdot\frac{\lambda^k}{k!}$	$F(k)=\sum_{i=0}^{k}e^{-\lambda}\cdot\frac{\lambda^i}{i!}$	λ	λ
Continuous				
Uniform(a,b)	$f(x)=\begin{cases}\frac{1}{b-a}, & a\le x\le b\\ 0, & \text{elders.}\end{cases}$	$F(x)=\begin{cases}0, & x<a\\ \frac{x-a}{b-a}, & a\le x<b\\ 1, & x\ge b\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentieel(λ)	$f(x)=\lambda e^{-\lambda x}$, $x\ge 0$	$F(x)=\begin{cases}1-e^{-\lambda x}, & x\ge 0\\ 0, & x<0\end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Veelgebruikte functies op de grafische rekenmachine

Kansverdeling van X	Type vraag	TI-84 Plus	Casio
Continu (willekeurig)	$P(a\le X\le b)$	$\int_a^b f(x)\,dx$	$\int_a^b f(x)\,dx$
Binomiaal(n,p)	$P(X=k)$ $P(X\le k)$	binompdf(n,p,k) binomcdf(n,p,k)	BinomialPD(k,n,p) BinomialCD(k,n,p)
N(μ,σ)	$P(a\le X\le b)$ Wat is g als $P(X\le g)=p$?	normalcdf(a,b,μ,σ) invNorm(opp,μ,σ)	normalCD(a,b,σ,μ) invNorm(opp,σ,μ)

- Continue kansvariabelen met kansdichtheidsfunctie f:

$$\int_a^b f(x)\,dx = \text{fnInt}(f(x);x;a;b)$$

- Binomiaal verdeelde kansvariabele $X\sim$ Binomiaal$(n;p)$:
  - $P(X=k)=$ binompdf$(n;p;k)$ en $P(X\le k)=$ binomcdf$(n;p;k)$
- Normaal verdeelde kansvariabele $X\sim N(\mu;\sigma)$:
  - $P(a\le X\le b)=$ normalcdf$(a;b;\mu;\sigma)$ en $P(X\le g)=p\to g=$ invNorm$(p;\mu;\sigma)$
- Poisson verdeelde kansvariabele $X\sim$ Poisson(λ):
  - $P(X=k)=$ poissonpdf$(\lambda;k)$ en $P(X\le k)=$ poissoncdf$(\lambda;k)$

z-score:

$$z=\frac{x-\mu}{\sigma}$$

Centrale limietstelling: Gegeven n kansvariabelen $X_1,X_2,\ldots,X_n$ die onderling onafhankelijk zijn en dezelfde kansverdeling hebben met een verwachtingswaarde μ en standaardafwijking σ, dan geldt (bij benadering) dat

- de som $\sum X = X_1 + X_2 + \ldots + X_n$ normaal verdeeld is met verwachtingswaarde $n \cdot \mu$ en standaardafwijking $\sqrt{n} \cdot \sigma$.
- het gemiddelde $\overline{X} = \frac{X_1+X_2+\ldots+X_n}{n}$ normaal verdeeld is met verwachtingswaarde μ en standaardafwijking $\frac{\sigma}{\sqrt{n}}$.