

# Formuleblad Statistiek deel 1 (2024-2025)

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

**Steekproefgemiddelde:**

$$\bar{x} = \frac{\sum_i x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

**Steekproefvariantie:**

$$s^2 = \frac{\sum_i (x_i - \bar{x})^2}{n} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n} \quad (\text{optie 1})$$

$$s^2 = \frac{\left(\sum_i x_i^2\right) - n \cdot \bar{x}^2}{n} = \frac{(x_1^2 + x_2^2 + \dots + x_n^2) - n \cdot \bar{x}^2}{n} \quad (\text{optie 2})$$

**Rekenregels kansrekening:**

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

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

$$P(A | B) = \frac{P(A \text{ en } B)}{P(B)} \quad (\text{conditionele kansen})$$

**Discrete en continue kansverdelingen:**

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

**Verwachtingswaarde en variantie van veelgebruikte kansverdelingen:**

Verdeling	Kans(dichtheids)functie	CDF	$E(X)$	$\text{Var}(X)$
<b>Discreet</b>				
Uniform( $a, b$ )	$p(k) = \frac{1}{b-a+1}$ ( $k = a, a+1, \dots, b$ )	$F(k) = \begin{cases} 0 & x < a \\ \frac{k-a+1}{b-a+1} & a \leq k < b \\ 1 & k \geq b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a+1)^2-1}{12}$
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( $\lambda$ )	$p(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$	$F(k) = \sum_{i=0}^k e^{-\lambda} \cdot \frac{\lambda^i}{i!}$	$\lambda$	$\lambda$
<b>Continuous</b>				
Uniform( $a, b$ )	$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{elders.} \end{cases}$	$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & x \geq b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentieel( $\lambda$ )	$f(x) = \lambda e^{-\lambda x}, x \geq 0$	$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

**z-score:**

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

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

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