## Probability

$P(A) \geq 0$

$P(\Omega) = 1$

$P(A)$ — probability

event $(A \subseteq \Omega)$

$A \subseteq \Omega$, $\Omega$ — event, event space

$\Omega \setminus A$ : the event $A$ does NOT happen

$A \cap B$ : both events happen

$A \cup B$ : either $A$ or $B$ happens

$P(A \cup B) = P(A) + P(B)$

$A \cap B = \emptyset$

**Conditional probability**

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

(event: $A$, observation: $B$)

**The Bayes formula**

$$\underbrace{P(A \mid B)}_{\text{posterior}} = \frac{\overbrace{P(B \mid A)}^{\text{likelihood}}\ \overbrace{P(A)}^{\text{prior}}}{P(B)}$$

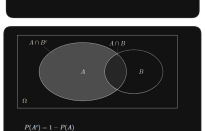
## Wetten morgan

$(A \cap B)^c = A^c \cup B^c$

$(A \cup B)^c = A^c \cap B^c$

Disjunct: $A \cap B = \emptyset$

$P(A \cup B) = P(A) + P(B)$




$P(A^c) = 1 - P(A)$

$P(A) = P(A \cap B) + P(A \cap B^c)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$P(A \cap B^c) = P(A) - P(A \cap B)$

## k-Permutaties

Volgorde maakt niet uit

$$P(n, k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

($n$ = aantal in verzameling, $k$ = aantal geselecteerd)

## k-Permutaties

De volgorde

$$P(n, k) = \frac{n!}{(n-k)!}$$

($n$ = aantal in verzameling, $k$ = aantal geselecteerd)

| Eigenschap | Permutatie $P(n, k)$ | Combinatie $\binom{n}{k}$ |
|---|---|---|
| Definitie | Geordende selectie van $k$ elementen uit $n$ | Niet-geordende selectie van $k$ elementen uit $n$ |
| Formule | $P(n, k) = \frac{n!}{(n-k)!}$ | $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ |
| Volgorde belangrijk? | Ja | Nee |
| Voorbeeld | Kies 3 winnaars voor de 1e, 2e en 3e plaats in een race | Kies 3 mensen voor een commissie |

$$n^{\underline{k}} = \frac{n!}{(n-k)!}$$

(Permutation of Full Set; Permutation of Selected Set)

$$n^{\underline{k}} = \frac{n!}{(n-k)!\,k!}$$

(Combination of Full Set; Permutation of Selected Set)

## PMF: Probalitietsmassafunctie

$p_X$: probability density function of $X$

$x$: possible values of $X$

$X$: probability of heads (random variable)

$$p_X(x) = \begin{cases} 1/6 & \text{if } x = 1 \\ 1/6 & \text{if } x = 2 \\ 1/6 & \text{if } x = 3 \\ 1/6 & \text{if } x = 4 \\ 1/6 & \text{if } x = 5 \\ 1/6 & \text{if } x = 6 \\ 0 & \text{otherwise} \end{cases}$$

Drager

## Marginale PMF

$$P(X = x) = \sum_y P(X = x, Y = y)$$

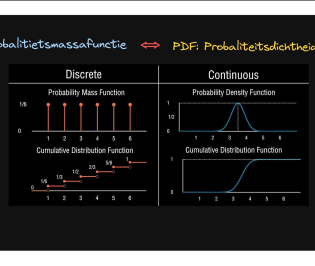
$$P(Y = y) = \sum_x P(X = x, Y = y)$$

| x\y | 1 | 2 | 3 |
|---|---|---|---|
| 1 | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |
| 2 | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |
| 3 | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

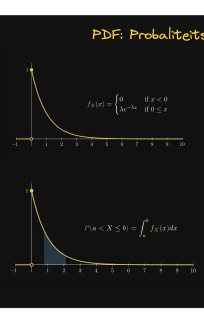
$P(X = x) = \sum_y \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$

$P(Y = y) = \sum_x \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$

## PMF: Probalitietsmassafunctie ⟺ PDF: Probaliteitsdichtheidsfunctie




## PDF: Probaliteitsdichtheidsfunctie




CDF: $F_X'(x) = f_X(x)$ (if the derivative exists)

PDF

**Rekeregels**

$P(a \leq X \leq b) = P(X \in [a, b]) = \int_a^b f_X(x)\,dx$

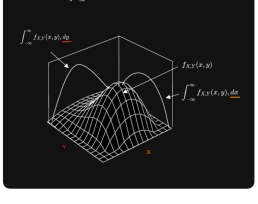
$P(a \leq X) = P(X \in [a, \infty)) = \int_a^\infty f_X(x)\,dx$

$P(X \leq b) = P(X \in ]-\infty, b]) = \int_{-\infty}^b f_X(x)\,dx$

## Marginale PDF

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dx$$



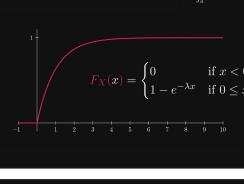

## CDF: Cumulative distributie functie

cumulative distribution function (CDF)

$$F_X(x) = P(X \leq x)$$

$$f_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ \lambda e^{-\lambda x} & \text{if } 0 \leq x \end{cases}$$

$$\int_a^b f(x)\,dx \rightarrow F_X(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 - e^{-\lambda x} & \text{if } 0 \leq x \end{cases}$$




**Onafhankelijke toevalsveranderlijken**

$$F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y)$$

**Uitwisselbare toevalsveranderlijken**

$$F_{X,Y}(x, y) = F_{X,Y}(y, x)$$

**Vermenigvuldigingsregel**

Discreet

$$p_{X,Y}(x, y) = p_{X|Y}(x \mid y) \cdot p_Y(y)$$

$$p_{X,Y}(x, y) = p_{Y|X}(y \mid x) \cdot p_X(x)$$

Continue

$$f_{X,Y}(x, y) = f_{X|Y}(x \mid y) \cdot f_Y(y)$$

$$f_{X,Y}(x, y) = f_{Y|X}(y \mid x) \cdot f_X(x)$$

**Conditionele verdeling**

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} \quad \text{voor } f_Y(y) > 0$$

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x, y)}{f_X(x)} \quad \text{voor } f_X(x) > 0$$

**Regel Bayes**

$$f_{X|Y}(x \mid y) = \frac{f_{Y|X}(y \mid x) \cdot f_X(x)}{f_Y(y)}$$

**Regel Totale probaliteit**

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y|X}(y \mid x) \cdot f_X(x)\,dx$$

## Disjunct ⟺ onafhankelijk

| Eigenschap | Disjuncte gebeurtenissen | Onafhankelijke gebeurtenissen |
|---|---|---|
| Definitie | $P(A \cap B) = 0$ | $P(A \cap B) = P(A)P(B)$ |
| Kunnen ze tegelijk optreden? | Nee, $A$ en $B$ sluiten elkaar uit | Ja, ze kunnen tegelijk optreden |
| Rekenkundige regel | $P(A \cup B) = P(A) + P(B)$ (voor disjuncte gebeurtenissen) | $P(A \cap B) = P(A)P(B)$ |
| Beïnvloeden ze elkaar? | Ja, als $A$ optreedt, kan $B$ niet optreden | Nee, de kans op $A$ verandert niet als $B$ optreedt, en vice versa |

## Conditioneel onafhankelijk

$P(A \cap B|C) = P(A|C)P(B|C)$

$P(A|B \cap C) = P(A|C)$ als $P(B \cap C) > 0$

$P(B|A \cap C) = P(B|C)$ als $P(A \cap C) > 0$

## Totale probabiliteit

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i) \cdot P(B_i)$$

## Verwachtingswaarde

Gemiddelde van oneindig veel uitkomsten

average outcome on the long run

$$\mathbb{E}|X| = \sum_{k=1}^{\infty} x_k P(X = x_k)$$

(outcome, probability of the outcome)

$$\mathbb{E}|X| = \sum_{\text{outcomes}} (\text{outcome}) \times P(\text{outcome})$$

PMF $\rightarrow E[X] = \sum_i x_i P(X = x_i)$

PDF $\rightarrow E[X] = \int_{-\infty}^{\infty} x f(x)\,dx$

## Eigenschappen

**Regel van verwachtingswaarde**

$E[g(X)] = \sum_x g(x) P(X = x)$ (discreet)

$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\,dx$ (continu)

**Totale verwachting**

$$E[Y] = E[E[Y \mid X]]$$

**Onafhankelijke toeval**

$$E[X \cdot Y] = E[X] \cdot E[Y]$$

**Lineare functie toevalsveranderlijke**

$$E[aX + b] = aE[X] + b$$

$$E[aX + bY] = a \cdot E[X] + b \cdot E[Y]$$

**Conditionel**

$E[X \mid Y] = \sum_x x P(X = x \mid Y)$ (discreet)

$E[X \mid Y] = \int x f_{X|Y}(x \mid Y)\,dx$ (continu)

**Iteratief**

$$E[X] = E[E[X \mid Y]]$$

PMF: $E[X] = \sum_y E[X \mid Y = y] P(Y = y)$

PDF: $E[X] = \int E[X \mid Y = y] f_Y(y)\,dy$

## Vector

**Discreet**

$$E[g(X, Y)] = \sum_x \sum_y g(x, y) \cdot p_{X,Y}(x, y)$$

**Continue**

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \cdot f_{X,Y}(x, y)\,dx\,dy$$

## Variantie en staardafwijking

De spreiding

| Maatstaf | Formule |
|---|---|
| Variantie | $\mathrm{Var}(X) = E[(X - E[X])^2]$ |
| Alternatieve vorm | $\mathrm{Var}(X) = E[X^2] - (E[X])^2$ |
| Standaardafwijking | $\sigma_X = \sqrt{\mathrm{Var}(X)}$ |

## Eigenschappen

**Lineare Functie**

$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$$

**Conditionele variantie**

$$\mathrm{Var}(X \mid Y) = E\left[(X - E[X \mid Y])^2 \mid Y\right]$$

**Totale variantie**

$$\mathrm{Var}(X) = E[\mathrm{Var}(X \mid Y)] + \mathrm{Var}(E[X \mid Y])$$

## Variantiesomregel

**X en Y onafhankelijk**

$$\mathrm{Var}(aX + bY) = a^2 \cdot \mathrm{Var}(X) + b^2 \cdot \mathrm{Var}(Y)$$

**X en Y afhankelijk**

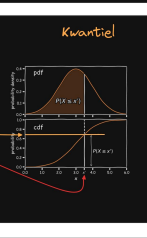
$$\mathrm{Var}(aX + bY) = a^2 \cdot \mathrm{Var}(X) + b^2 \cdot \mathrm{Var}(Y) + 2ab \cdot \mathrm{Cov}(X, Y)$$

$$\mathrm{Var}(X + Y + Z) = \mathrm{Var}(X) + \mathrm{Var}(Y) + \mathrm{Var}(Z) + 2\mathrm{Cov}(X, Y) + 2\mathrm{Cov}(X, Z) + 2\mathrm{Cov}(Y, Z)$$

## Kwantiel

Inverse CDF

$$Q(p) = F_X^{-1}(p)$$




**Interkwantielafstand**

$$\mathrm{IQR} = F_X^{-1}(0.75) - F_X^{-1}(0.25)$$

## Moment

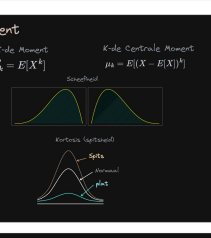
| Moment | Definitie | Betekenis |
|---|---|---|
| 1e moment ($\mu_1'$) | $E[X]$ | Gemiddelde |
| 2e moment ($\mu_2'$) | $E[X^2]$ | Wordt gebruikt voor variantie |
| 2e centrale moment ($\mu_2$) | $E[(X - \mu)^2]$ | Variant |
| 3e centrale moment ($\mu_3$) | $E[(X - \mu)^3]$ | Scheefheid (asymmetrie) |
| 4e centrale moment ($\mu_4$) | $E[(X - \mu)^4]$ | Kurtosis (spitsheid) |

K-de Moment

$$\mu_k' = E[X^k]$$

K-de Centrale Moment

$$\mu_k = E[(X - E[X])^k]$$




## Kwadratische fout

**Mean Squared Error (MSE)**

Kwadratisch grote fouten zwaarder

$$\mathrm{MSE} = E\left[(Y - \hat{Y})^2\right]$$

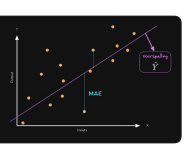
**Mean Absolute Error (MAE)**

Minder gevoelig

$$\mathrm{MAE} = E\left[|Y - \hat{Y}|\right]$$

## Kleinste-kwadraten-schatter

| Kleinste-kwadraten-schatter | Formule |
|---|---|
| Algemene schatter | $\hat{X}(Y) = \mathrm{E}(X \mid Y)$ |
| Lineaire schatter | $\hat{X}(Y) = \mathrm{E}(X) + \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}(Y - \mathrm{E}(Y))$ |




## Covariantie - correlatie

Linear verband tussen twee toevalsveranderlijken

$$\mathrm{Cov}(X, Y) = E\left[(X - E[X])(Y - E[Y])\right]$$

**Lineaire schaling**

$$\mathrm{Cov}(aX + b, Y) = a \cdot \mathrm{Cov}(X, Y)$$

**Verwachtingswaarde-eigenschap**

$$\mathrm{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

**Lineaire combinatie**

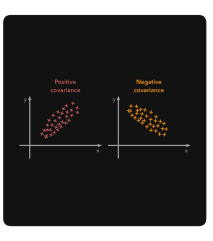
$$\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$$

**Correlatiecoëfficiënt**

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

$\sigma_X = \sqrt{\mathrm{Var}(X)}$

$\sigma_Y = \sqrt{\mathrm{Var}(Y)}$

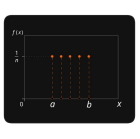



## Discrete verdelingen

## Continue verdelingen

## Discrete Uniforme Verdeling

Alle mogelijkheden hebben een gelijke kans




| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathrm{UniD}(a, b)$ |
| Drager | $x \in a, a+1, \ldots, b$ |
| PMF | $P(X = x) = \frac{1}{b-a+1}$ |
| Gemiddelde | $E[X] = \frac{a+b}{2}$ |
| Mediaan | $\frac{a+b}{2}$ |
| Variantie | $\frac{(b-a)(b-a+2)}{12}$ |

## Bernoulli verdeling

experiment met twee mogelijke uitkomsten: succes (1) met kans $p$ en mislukking (0) met kans $1-p$

| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathrm{Ber}(p)$ |
| Drager | $x \in 0, 1$ |
| PMF | $P(X = x) = p^x(1-p)^{1-x}$ |
| Gemiddelde | $p$ |
| Mediaan | 0 als $p < 0.5$, anders 1 |
| Variantie | $p(1-p)$ |

## Binomiale Verdeling

aantal successen in $n$ onafhankelijke Bernoulli-experimenten met kans $p$

| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathrm{Bin}(n, p)$ |
| Drager | $x \in 0, 1, \ldots, n$ |
| PMF | $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ |
| Gemiddelde | $np$ |
| Mediaan | $\lfloor np \rfloor$ |
| Variantie | $np(1-p)$ |

## Negatieve Binomiale verdeling

Aantal Bernoulli-experimenten nodig zijn om $r$ successen te behalen

| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathrm{NegBin}(r, p)$ |
| Drager | $x \in r, r+1, r+2, \ldots$ |
| PMF | $P(X = k) = \binom{k+r-1}{k} p^r (1-p)^k$ |
| Gemiddelde | $\frac{r(1-p)}{p}$ |
| Mediaan | / |
| Variantie | $\frac{r(1-p)}{p^2}$ |

## Geometische verdeling

Aantal Bernoulli-proeven tot de eerste succes

| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathrm{Geo}(p)$ |
| Drager | $x \in 1, 2, 3, \ldots$ |
| PMF | $P(X = k) = (1-p)^{k-1} p$ |
| Gemiddelde | $\frac{1-p}{p}$ |
| Mediaan | $\left\lceil -\frac{\ln(2)}{\ln(1-p)} \right\rceil$ |
| Variantie | $\frac{1-p}{p^2}$ |

## Poisson verdeling

Modelleert het aantal gebeurtenissen in een vast tijdsinterval.

| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathrm{Poi}(\lambda)$ |
| Drager | $x \in 0, 1, 2, \ldots$ |
| PMF | $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ |
| Gemiddelde | $\lambda$ |
| Mediaan | $\approx \lfloor \lambda \rfloor$ |
| Variantie | $\lambda$ |

## Hypergeometrische Verdeling

aantal successen in een steekproef zonder teruglegging met groepen A en B

| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathrm{Hypergeo}(A, B, n)$ |
| Drager | $x \in 0, \ldots, n$ |
| PMF | $P(X = x) = \frac{\binom{A}{x}\binom{B}{n-x}}{\binom{A+B}{n}}$ |
| Gemiddelde | $n\frac{A}{A+B}$ |
| Mediaan | - |
| Variantie | $\frac{nAB}{(A+B)^2}\frac{A+B-n}{A+B-1}$ |

## Exponentiële Verdeling

Tijd tussen twee Poisson-gebeurtenissen

| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathrm{Exp}(\lambda)$ |
| Drager | $x \geq 0$ |
| PDF | $f(x) = \lambda e^{-\lambda x}$ |
| Gemiddelde | $\frac{1}{\lambda}$ |
| Mediaan | $\frac{\ln(2)}{\lambda}$ |
| Variantie | $\frac{1}{\lambda^2}$ |

## Continue Uniforme Verdeling

Kans gelijk voor elke waarde binnen een interval $[a, b]$

| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathrm{UniC}(a, b)$ |
| Drager | $x \in [a, b]$ |
| PDF | $f(x) = \frac{1}{b-a}$ |
| Gemiddelde | $\frac{a+b}{2}$ |
| Mediaan | $\frac{a+b}{2}$ |
| Variantie | $\frac{(b-a)^2}{12}$ |

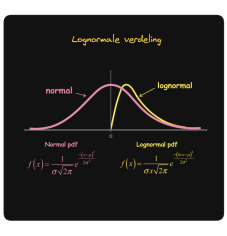
## Normale Verdeling

continue variabelen met een symmetrische verdeling

| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathcal{N}(\mu, \sigma^2)$ |
| Drager | $x \in \mathbb{R}$ |
| PDF | $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ |
| Gemiddelde | $\mu$ |
| Mediaan | $\mu$ |
| Variantie | $\sigma^2$ |

## Bivariaat normale verdeling

## Lognormale verdeling




Normal pdf

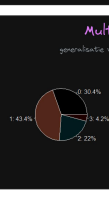
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Lognormal pdf

$$f(x) = \frac{1}{\sigma x\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}$$

## Multinomiale verdelingen (discreet)

generalisatie van de binomiale verdeling naar meer dan twee mogelijke uitkomsten




| Eigenschap | Waarde |
|---|---|
| Notatie | $\mathrm{MultBin}(n, p_1, \ldots, p_k)$ |
| Drager | $x_i \geq 0$ met $\sum x_i = n$ |
| PMF | $P(X_1 = x_1, \ldots, X_k = x_k) = \frac{n!}{x_1! x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$ |
| Gemiddelde | $E[X_i] = np_i$ |
| Mediaan | / |
| Variantie | $np_i(1 - p_i)$ |

## SISO discreet lineair

$$Y = aX + b \quad \text{met} \quad a \neq 0$$

$$p_Y(y) = p_X\left(\frac{y - b}{a}\right)$$

## SISO model

Single input single output

**Discreet**

$$p_Y(y) = \sum_{x|g(x)=y} p_X(x) = \sum_{g(x)=y} p_X(x)$$

Drager: $\mathrm{supp}(Y) = g(x) \mid x \in \mathrm{supp}(X)$

**Continue**

1) CDF bepalen

$$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(g(X) \leq y) = \int_{g(x) \leq y} f_X(x)\,dx$$

2) Afleiden naar PDF

$$f_Y(y) = \frac{dF_Y(y)}{dy}$$

($h(y)$, $h = g^{-1}$)

## SISO continue 1 stap

$$f_Y(y) = \left\{ f_X(h(y)) \left| \frac{dh}{dy}(y) \right| \right., \quad \text{als } y \in T\ 0, \quad \text{anders}$$

($h = g^{-1}$; Afleiden inverse $h(x)$)

**Lineare functie** (continue verdeling)

$$Y = aX + b \text{ met } a \neq 0$$

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y - b}{a}\right)$$

**Lineare functie** (Normaal verdeling)

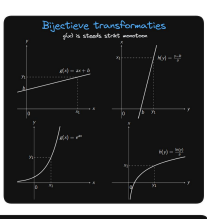
$$Y = aX + b \quad X \sim N(\mu, \sigma^2)$$

$$Y = \mathcal{N}(a\mu + b, a^2\sigma^2)$$

$\mathrm{Var}(Y) = a^2\sigma^2 \quad E[Y] = a\mu + b$

## Bijectieve transformaties

$g(x)$ is steeds strikt stijgend




## Twee discrete toevalsvectoren

$$Z = X + Y$$

$$p_Z(z) = \sum_{x+y=z} p_{X,Y}(x, y) = \sum_x p_{X,Y}(x, z - x)$$

## MIMO model

Multi input multi output

$\mathbf{X} = (X_1, X_2, \ldots, X_n)$ $\quad$ $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_m)$

Model: $(\mathbf{Y} = g(\mathbf{X}) = (g_1(\mathbf{X}), g_2(\mathbf{X}), \ldots, g_m(\mathbf{X}))$

**Discreet**

$$p_Y(\mathbf{y}) = \sum_{\mathbf{x}|g(\mathbf{x})=\mathbf{y}} p_X(\mathbf{x})$$

**Continue**

$$f_Y(\mathbf{y}) = f_X(\mathbf{h}(\mathbf{y})) \cdot |\det(J_h(\mathbf{y}))|$$

$\mathbf{h} = \mathbf{g}^{-1}$

$$J_h(y) = \begin{bmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} & \cdots & \frac{\partial h_1}{\partial y_n} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} & \cdots & \frac{\partial h_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_n}{\partial y_1} & \frac{\partial h_n}{\partial y_2} & \cdots & \frac{\partial h_n}{\partial y_n} \end{bmatrix}$$

## Probabiliteitsintegraaltransformatie (PIT)

Elke continue verdeling kan worden getransformeerd naar een andere verdeling via de PIT

Discrete CDF $F_X$

Continue CDF $G$

$$Z = G^{-1}(F_X(X))$$

$$Y = F_X(X) \Rightarrow Y \stackrel{d}{=} \mathrm{UniC}(0, 1)$$

$$Z = G^{-1}(Y) = G^{-1}(F_X(X)) \Rightarrow F_Z = G$$

## MISO model

**Convolutie integraal**

$$Y = X_1 + X_2$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(z) f_{X_2}(y - z)\,dz$$

$$X_1 = z, X_2 = y - z$$

| Verdeling van $X_i$ | Som $S = \sum_{i=1}^{k} X_i$ |
|---|---|
| $X_i \stackrel{d}{=} \mathrm{Bin}(n, p)$ | $S \stackrel{d}{=} \mathrm{Bin}(kn, p)$ |
| $X_i \stackrel{d}{=} \mathrm{Poi}(\lambda)$ | $S \stackrel{d}{=} \mathrm{Poi}(k\lambda)$ |
| $X_i \stackrel{d}{=} \mathrm{NegBin}(r, p)$ | $S \stackrel{d}{=} \mathrm{NegBin}(kr, p)$ |
| $X_i \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$ | $S \stackrel{d}{=} \mathcal{N}(k\mu, k\sigma^2)$ |

## CDF ordestatistiek

CDF van de k-de statistiek

$$F_{X_{(i)}}(x) = \mathbb{P}(X_{(i)} \leq x) = \sum_{j=i}^{n} \binom{n}{j} F(x)^j (1 - F(x))^{n-j}$$

## Steekproef

**Midden steekproefbereik**

$$\text{Midrange} = \frac{X_{(1)} + X_{(n)}}{2}$$

**Steekproefbereik**

$$\text{Bereik} = X_{(n)} - X_{(1)}$$

**Steekproefmediaan**

oneven $= X_{\left(\frac{n+1}{2}\right)}$

even $= \frac{1}{2}\left(X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}\right)$

**steekproefgemiddelde**

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$E[\bar{X}_n] = \mu$$

$$\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$$

## Markov-ongelijkheid

Als een niet-negatieve toevalsveranderlijke een kleine verwachtingswaarde heeft, dan is de kans op een grote waarde klein.

$$\mathrm{P}(X \geq a) \leq \frac{\mathrm{E}[X]}{a}$$

$X \geq 0$

$\mathrm{E}[X] < \infty$

## Chebyshev-ongelijkheid

Het steekproefgemiddelde heeft dezelfde verwachtingswaarde als de populatie, maar een kleinere spreiding rond zichzelf = Trouwens.

$\mu$: verwachtingswaarde (gemiddelde)

$\sigma^2$: variantie van $X$

$c$: een positieve constante afstand

$$\mathrm{P}(|X - \mu| \geq c) \leq \frac{\sigma^2}{c^2}$$

$$\mathrm{P}(|X - \mu| < c) \geq 1 - \frac{\sigma^2}{c^2}$$

$\mathrm{E}[X] = \mu$

$\mathrm{Var}(X) = \sigma^2 < \infty$

## Wet van de Grote Aantallen

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu \qquad \mathrm{E}[X_i] = \mu$$

**Centrale limietstelling**

$$S_n = \sum_{i=1}^{n} X_i \quad \bar{X}_n = \frac{S_n}{n}$$

$$\boxed{\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)}$$

$S_n \approx \mathcal{N}(n\mu, n\sigma^2)$

$\mathrm{E}[X_i] = \mu$

$$\boxed{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1)}$$

$\mathrm{Var}(X_i) = \sigma^2$

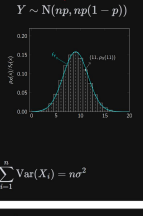
steekproefgemiddelde

**Continuiteitscorrectie**

Correctie voor discrete verdeling (binomiaal) $\quad \mathrm{Bin}(n, p)$

$$\mathrm{P}(a \leq X \leq b) \approx \mathrm{P}(a - 0.5 \leq Y \leq b + 0.5)$$

$$Y \sim \mathrm{N}(np, np(1 - p))$$




$$\mathrm{E}[S_n] = \sum_{i=1}^{n} \mathrm{E}[X_i] = n\mu \qquad \mathrm{Var}(S_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = n\sigma^2$$

## Poisson-benadering van de binomiale verdeling

Wanneer $n$ groot is en $p$ klein !

$$\lambda = np$$

$$\mathrm{Bin}(n, p) \approx \mathrm{Poisson}(\lambda)$$




## Chi-kwadraatverdeling

Chi-kwadraatverdeling met k vrijheidsgraden

$$Z_i = \mathcal{N}(0, 1)$$

$$\mathrm{E}[\chi_k^2] = k$$

$$\mathrm{Var}(\chi_k^2) = 2k$$

$$\boxed{\chi_k^2 = \sum_{i=1}^{k} Z_i^2}$$

$$\chi_k^2 \approx \mathcal{N}(k, 2k)$$

$$f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}$$

| Grootheid | Symbool/notatie |
|---|---|
| Notatie | $\chi_k^2$ |
| Parameter(s) | $k \in \mathbb{N}$ |
| Drager | $x \in [0, \infty)$ |
| PDF | $f(x)$ |
| Gemiddelde | $\mathrm{E}[X]$ |
| Variantie | $\mathrm{Var}(X)$ |

## Markov-keten

Elke rij zoek tot 1 $\sum_j P_{ij} = 1$

**transitiegraaf**




**transitiematrix**

$$T = \begin{array}{c|cc} \text{van} \backslash \text{naar} & & \\ \hline & 0.8 & 0.2 \\ & 0.6 & 0.4 \end{array} \qquad T = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \cdots & P_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ P_{m1} & P_{m2} & \cdots & P_{mm} \end{bmatrix}$$

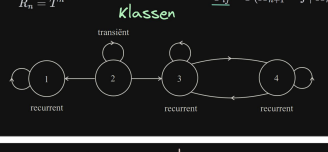
**Transitieprobabiliteiten**

$$P_{ij} = \mathrm{P}(X_{n+1} = j \mid X_n = i)$$

**n staps transtie**

$$R_n = T^n$$

## Klassen




## Chapman-Kolmogorov

Analyseren van lange-termijn gedrag, zoals stationaire verdelingen

$$R_n = T^n$$

$$r_{ij} = \mathrm{P}(X_n = j \mid X_0 = i)$$

$p_{kj}$: kans op overgang van $k$ naar $j$

$r_{ik}(n - 1)$: kans om in $k$ te zijn op tijd $n - 1$

$r_{ij}(n)$: kans dat men in toestand $j$ is op tijd $n$

$$r_{ij}(n) = \sum_{k \in j} r_{ik}(n - 1) \cdot p_{kj}$$

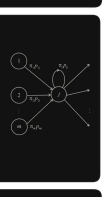
## Steady-state convergentie

Bij 1 recurrente aperiodieke klasse kent een keten steady state gedrag

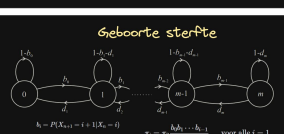
$$\lim_{n \to \infty} r_{ij}(n) = \pi_j \qquad \sum_{k=1}^{n} \pi_k = 1 \qquad \pi_j = \sum_{k=1}^{n} \pi_k \cdot P_{kj}$$

**Als matrix**

$$\begin{bmatrix} P_{00} - 1 & P_{10} & P_{20} \\ P_{01} & P_{11} - 1 & P_{21} \\ P_{02} & P_{12} & P_{22} - 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \pi_0 \\ \pi_1 \\ \pi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$




## Geboorte sterfte

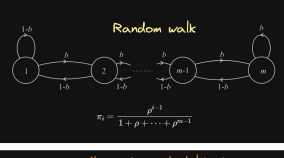



$b_i = P(X_{n+1} = i + 1 | X_n = i)$

$d_i = P(X_{n+1} = i - 1 | X_n = i)$

$$\pi_i = \pi_0 \frac{b_0 b_1 \cdots b_{i-1}}{d_1 d_2 \cdots d_i}, \quad \text{voor alle } i = 1, \ldots, m$$

## Random walk




$$\pi_i = \frac{\rho^{i-1}}{1 + \rho + \cdots + \rho^{m-1}}$$

## Absorptieprobabiliteiten

n toestanden en toestand s is absorberend

$a_s = 1$

$a_i =