H3: belangrijke verdelingen

3.1 eenvoudige univariabele verdelingen

3.1.1 uniforme verdeling

def: uniforme verdeling	= beschrijft de waarschijnlijkheid die constant is binnen bepaald gebied > a en b zijn grenzen van gebied > rechthoek: $P(x;a,b) = \begin{cases} \frac{1}{b-a} & \text{voor } a \leq x \leq b \\ 0 & \text{elders} \end{cases}$ > standaardafwijking hiervoor is de breedte gedeeld door √12 $V(X) = <X^2> - <X>^2 = \int_a^b x^2 \frac{1}{b-a} dx - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$
eigenschappen vd uniforme verdeling	- verwachtingswaarde: $<X> = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2}$ - variantie: $V(X) = <X^2> - <X>^2 = \int_a^b x^2 \frac{1}{b-a} dx - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$ > standaardafwijking is breedte gedeeld door √12 - cumulatieve verdelingsfunctie: $F(x;a,b) = \begin{cases} 1 & x > b \\ \frac{x-a}{b-a} & \text{voor } a \leq x \leq b \\ 0 & x < a \end{cases}$

3.1.2 exponentiële verdeling

def: exponentiële verdeling	Voor een variabele X (0≤x≤∞) is deze: $P(x;\xi) = \frac{1}{\xi} e^{-x/\xi}$ > gekarakteriseerd door één parameter ξ
eigenschappen exponentiële verdeling	- verwachtingswaarde: $<X> = \frac{1}{\xi} \int_0^\infty x e^{-x/\xi} dx = \xi$ - variantie: $V(X) = \frac{1}{\xi} \int_0^\infty (x-\xi)^2 e^{-x/\xi} dx = \xi^2$ - cumulatieve verdelingsfunctie: $F(x;\xi) = \begin{cases} 1 - e^{-x/\xi} & x \geq 0 \\ 0 & x < 0 \end{cases}$
stelling: geheugenloos	De exponentiële functie is geheugenloos: $P(X > x_1 + x_2 \| X > x_2) = P(X > x_1)$

3.2 binomiale verdeling

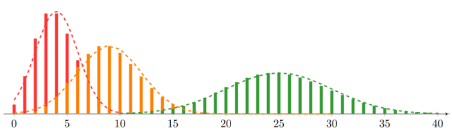
3.2.1 Bernoulli proces

def: Bernoulli verdeeld	Een toevallige variabele x is Bernoulli verdeeld als: > X slechts twee waarden 0 of 1 kan aannemen met kansen 1-p en p > dan: $P(x;p) = p^x (1-p)^{1-x}$ met 0≤p≤1 de kans van succes
eigenschappen Bernoulli verdeling	- verwachtingswaarde: <X> = p - variantie: V(X) = p(1-p)

3.2.2 Binomiale verdeling

binomiale verdeling	= n opeenvolgende Bernoulli-exp. met zelfde succeskans p die onafhankelijk van elkaar uitgevoerd worden > bin. verd. geeft de waarschijnlijkheid om x keer succes te hebben $P(x;p,n) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$

ontleding van binomiale verdeling	1: permutaties 2: successen
1: permutaties	er zijn 2^n mogelijke permutaties van succes of falen > aantal met x successen = aantal manieren om x te selecteren in n: $$C_x^n = \binom{n}{x} = \frac{n!}{x!(n-x)!}$$
2: successen	er zijn x successen met waarschijnlijkheid p x-1 1-p > totale waarschijnlijkheid bekomen door ze te vermenigvuldigen: $p^x(1-p)^{n-x}$.
stelling: verwachtingswaarde bin.verd 3.2.2	de verwachtingswaarde ve toev. verdeelde X geeft het gemiddeld aantal successen $$< X >= np$$
stelling: variantie bin. verd 3.2.3	de variantie ve bin. verdeelde toev. variabele wordt gegeven door: $$V(X) = np(1-p) \qquad \sigma = \sqrt{np(1-p)}$$
3.3 Poisson verdeling	
doel Poisson verdeling	Poissonverdeling beschrijft exp. waarbij je een bepaald aantal uitkomsten hebt, maar niet weet hoeveel pogingen > individuele pogingen in een continuüm
3.3.1 waarschijnlijkheidsverdeling	
def: Poissonverdeling	= beschrijft processen met bepaalde uitkomsten X, maar onbepaald #pogingen > we def. λ = gemiddeld aantal evenementen die verwacht worden > waarschijnlijkheid om x-evenementen waar te nemen is: $$P(x;\lambda) = \frac{e^{-\lambda}\,\lambda^x}{x!}$$
> afleiden Poissonverdeling	veronderstel er wordt gemiddeld λ evenementen verwacht > splits interval op in n zeer kleine intervalletjes > waarschijnlijkheid om in een intervalletje te liggen is: p=λ/n Waarschijnlijkheid dat er X evenementen in n deelintervalletjes van het grote interval zullen optreden is gegeven door binomiaal: $$P(x;n,\frac{\lambda}{n}) = \frac{n!}{x!(n-x)!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$ voor n→∞ met x eindig vinden we: $$\lim_{n\to\infty}\frac{n!}{(n-x)!} = \lim_{n\to\infty} n(n-1)(n-2)\cdots(n-x+1) = n^x$$ en er verschijnt een exponentiële : $$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n-x} = \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$$ > deze geeft dan de Poissonverdeling
stelling: verwachtingswaarde P.verd 3.3.1	<X> = λ
stelling: variantie en σ P.verd 3.3.2	$V(X) = \lambda \qquad \sigma_X = \sqrt{\lambda}$

3.3.2 twee Poisson verdelingen	
stelling: 2 Poissonverdelingen	Zei er twee types evenementen a en b Zei er twee P.verd voor deze a en b met λ_a en λ_b > waarschijnlijkheid voor X evenementen is ook Poisson verdeelt met: $\lambda = \lambda_a + \lambda_b$
stelling: meerdere P.verdelingen	Voor een bepaald aantal evenementen x is de verdeling ch aantal evenementen van type a gegeven door een binomiale distributie $P_{\text{bin}}(x_a; x, p_a).$
3.4 Normale (Gauss) verdeling	
Gauss verdeling	$P(X;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}$ > - middelpunt rond X=μ - breedte bepaald door σ - bij X=μ+-σ is P(X) = 61%
def: eenheidsgauss	Substitueer: $Z = \frac{(X-\mu)}{\sigma}$ dan wordt de Gauss $\mathcal{N}(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$
normalisatie van de gauss	er geldt: $\int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{2\pi}\sigma$ dus de Gauss is genormaliseerd op 1
verwachtingswaarde Gauss	<X> = μ
variantie Gauss	$\langle (x-\mu)^2 \rangle = \sigma^2$
3.4.2 cumulatieve verdeling en errorfunctie	
normale cumulatieve verdelingsfunctie Gauss	Kan niet analytisch berekend worden > opzoeken in tabellen
def: Errorfunctie	= waarschijnlijkheidsdichtheid vd eenheidsgauss > normale verdeling met μ=0 en σ=1 $\mathrm{Erf}(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} du$
Gauss als errorfunctie	Elke Gauss verdeling kan teruggevonden worden door substitutie in de errorftie $X = \sigma Z + \mu.$ > c.d.f van deze Gauss kan worden bepaald door de errorfunctie: $\mathrm{Erf}\left(\frac{X-\mu}{\sigma}\right)$
enkele handige waarden	• 68.27% van de oppervlakte ligt binnen 1σ van het gemiddelde, • 95.45% ligt binnen 2σ, • 99.73% ligt binnen 3σ. Als je ronde getallen wil in de percentages, dan: • 90% ligt binnen 1.645σ, • 95% ligt binnen 1.960σ, • 99% ligt binnen 2.576σ, • 99.9% ligt binnen 3.290σ.
3.4.3 Gauss als limiet van Poisson en binomiaal	
verband Gauss-Poisson	voor grote λ lijkt de Poisson verdeling op een Gauss met μ=λ en σ=√λ > kan vanaf λ=10 
verband Gauss-binomiaal	binomiaal neigt naar een Gauss met μ=np en $\sigma = \sqrt{np(1-p)}.$ > rond p=0.5

3.4.4 lognormale verdeling	
def: lognormale verdeling	een toevallige variabele X is lognormaal verdeelt indien de waarschijnlijkheids-dichtheid gegeven wordt door: $f(X;\mu,\sigma) = \frac{1}{\sqrt{2\pi}x\sigma} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}$ > met x>0
eigenschappen lognormale verdeling	- verwachtingswaarde: $< X >= e^{\frac{1}{2}(2\mu+\sigma^2)}$ - variantie: $V(X) = \left(e^{\sigma^2} - 1\right) e^{(2\mu+\sigma^2)}$
stelling: verband lognormale verd.	Zei X een toev. variabele die lognormaal verdeelt is met parameter μ en σ > de getransformeerde toev. var. Y = ln(X) is ook normaal verdeeld met: - verwachtingswaarde = μ - variantie = σ
3.4.5 Breit-Wigner (Cauchy) distributie	
Breit-Wigner functie	voor deeltjes met massa m afkomstig van een resonantie M en breedte Γ: $F(m;M,\Gamma) = \frac{1}{2\pi}\frac{\Gamma}{(m-M)^2 + \left(\frac{\Gamma}{2}\right)^2}$ Γ is de FWHM
eigenschappen B-W functie	- ze kan omgezet worden in Cauchy verdeling door verandering van oorsprong en schaal $F(z) = \frac{1}{\pi(1+z^2)}$ - de integralen vd functie divergeren > geen momenten, verwachtingswaarde, variantie - wet van grote getallen is niet geldig
3.5 multi-dimensionale verdelingen	
3.5.1 multinomiale verdeling	
def: multinomiale verdeling	= beschrijft processen met n identieke onafh. pogingen van m mogelijke uitkomsten $X_{(1)},\ldots,X_{(m)}$ met individuele succeskans $p_1,\ldots,p_m$ $p(x_{(1)},\ldots,x_{(m)};n,p_1,\ldots,p_m) = \frac{n!}{x_{(1)}!\ldots x_{(m)}!}p_1^{x_{(1)}}\ldots p_m^{x_{(m)}}$ waarbij $x_{(1)} + x_{(2)} + \ldots + x_{(m)} = n$ en $p_1 + p_2 + \ldots + p_m = 1$ (multidimensionale binomiale verdeling)
3.5.2 multi-dimensionale Gauss verdeling	
def: multidimensionale Gauss	beschouw een verdeling van m variabelen in $X_{(1)},\ldots,X_{(m)}$ > compact genoteerd door **X** > gemiddelde $\mu_{(1)},\ldots,\mu_{(m)}$ worden beschreven door $\boldsymbol{\mu}$ Dan is de Gauss: $P(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \lvert\mathbf{V}\rvert}}\exp\left[-\frac{1}{2}(\tilde{\mathbf{x}} - \tilde{\boldsymbol{\mu}})\mathbf{V}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right]$

> afleiden multidimensionale Gauss	Meest algemene vorm voor de Gauss is een exponentiële ve kwadratische vorm die termen bevat in $X_{(i)}^2$, kruistermen in $X_{(i)}X_{(j)}$, lineaire termen en een constante > dus: $$P(\mathbf{x}) \propto \exp\left[-\frac{1}{2}(\tilde{\mathbf{x}} - \tilde{\boldsymbol{\mu}})\mathbf{A}(\mathbf{x} - \boldsymbol{\mu})\right]$$ Met deze notatie zijn $\mathbf{X}$ en $\boldsymbol{\mu}$ kolomvectoren. Tenslotte eisen we dat $\mathbf{A}$ symmetrisch is: $A_{ij} = A_{ji}$. 1: Bekijk het speciaal geval waarbij A diagonaal is > P(**X**) factoriseert in m onafhankelijke Gaussen: $$\exp\left[-\frac{1}{2}\sum_{i=1}^{m} A_{ii}X_{(i)}^2\right] = \prod_{i=1}^{m} e^{-\frac{A_{ii}}{2}X_{(i)}^2}$$ waarbij de diagonale elementen: $$A_{ii} = \frac{1}{\sigma_i^2}$$ aangezien A diagonaal is kunnen we: $$\mathbf{A} = \mathbf{V}^{-1}$$ wat ons geeft: $$P(\mathbf{x}) \propto \exp\left[-\frac{1}{2}\tilde{\mathbf{x}}\mathbf{V}^{-1}\mathbf{x}\right]$$ 2: bekijk nu een algemeen geval: > we kunnen A diagonaliseren: $$\mathbf{U}\mathbf{A}\tilde{\mathbf{U}} = \mathbf{A}'$$ De exponent $\tilde{\mathbf{X}}\mathbf{A}\mathbf{X}$ in vgl.3.21 kan geschreven worden als: $$\tilde{\mathbf{X}}\,(\tilde{\mathbf{U}}\mathbf{U})\,\mathbf{A}\,(\tilde{\mathbf{U}}\mathbf{U})\,\mathbf{X}$$ Dit is niets anders dan $\tilde{\mathbf{X}}'\mathbf{A}'\mathbf{X}'$ met $\mathbf{A}'$ diagonaal. De variantiematrix $\mathbf{V}'$ voor de $\mathbf{X}' = \mathbf{U}\mathbf{X}$ is dus diagonaal met elementen$(\mathbf{U}\mathbf{A}\tilde{\mathbf{U}})^{-1} = \mathbf{U}\mathbf{A}^{-1}\tilde{\mathbf{U}}$. We kennen dus de variantiematrix voor de $\mathbf{X}'$ en ook de relatie met de $\mathbf{X}$ via: $\mathbf{X} = \tilde{\mathbf{U}}\mathbf{x}'$. We gebruiken nu een formule die de variantiematrix geeft voor een set variabelen die een functie zijn van een andere set. Dit resultaat is niet-triviaal en zal besproken worden in het hoofdstuk "foutenpropagatie van functies van verschillende variabelen". (Merk op dat in dit speciale geval de relatie lineair is, en de vergelijking dus exact is en geen benadering.): $$\begin{aligned}\mathbf{V} &= \tilde{\mathbf{U}}\mathbf{V}'\mathbf{U} \\ &= \tilde{\mathbf{U}}\mathbf{U}\mathbf{A}^{-1}\tilde{\mathbf{U}}\mathbf{U} \\ &= \mathbf{A}^{-1}\end{aligned}$$
binomiale verdeling	Voor twee dimensies geldt: $$\mathbf{V} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$ met als inverse: $$\mathbf{V}^{-1} = \frac{1}{(1-\rho^2)\sigma_x^2\sigma_y^2}\begin{pmatrix} \sigma_y^2 & -\rho\sigma_x\sigma_y \\ -\rho\sigma_x\sigma_y & \sigma_x^2 \end{pmatrix}$$ De volledige formule (met inbegrip van de normering) voor de binormale of twee-dimensionale Gauss is dus: $$P(X,Y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{(1-\rho^2)}}\exp\left[-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right]\right]$$ In twee dimensies is de unitaire matrix U die de exponent diagonaliseert, de rotatiematrix: $$\mathbf{U} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$ > het X,Y systeem wordt gevonden door het U,V systeem te roteren over een hoek θ > we kunnen via algebra de relaties vinden: $$\rho\sigma_x\sigma_y = \cos\theta\sin\theta(\sigma_v^2 - \sigma_u^2) \qquad \tan 2\theta = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2}$$ $$\sigma_x^2 = \cos^2\theta\sigma_u^2 + \sin^2\theta\sigma_v^2 \qquad \sigma_u^2 = \frac{\cos^2\theta\sigma_x^2 - \sin^2\theta\sigma_y^2}{\cos^2\theta - \sin^2\theta}$$ $$\sigma_y^2 = \cos^2\theta\sigma_v^2 + \sin^2\theta\sigma_u^2 \qquad \sigma_v^2 = \frac{\cos^2\theta\sigma_y^2 - \sin^2\theta\sigma_x^2}{\cos^2\theta - \sin^2\theta}$$