Statistiek voor MBW/KW 2023-2024 Week 2

- Kansrekening
- Discrete kansvariabelen (stochasten), kansverdelingen
- Verwachtingswaarde, variantie, standaarddeviatie
- Rekenregels
- Bernoulliverdeling
- Binomiale verdeling
- Operation Eagle Claw

Kansrekening: Ontstaan van de kansrekening

Kansspelen zijn spellen waarbij het toeval een grote rol speelt, vaak met dobbelstenen, munten of kaarten.

Ze hebben een verslavende werking, vooral als er om geld gespeeld wordt.

Bekend uit de oudheid, Grieken, Romeinen. Niet alleen voor soldaten en het lage volk, maar ook gepractiseerd door de adel.

Een 17^{e} eeuws dobbelspel

Bron:https://carnotcycle.wordpress.com/tag/antoine-gombaud/

Kansrekening: De paradox van Chevalier de Méré

Antoine Gombaud, **Chevalier de Méré** (1607 – 1684), schrijver en verwoed gokker.

Beroemde citaten:

Bewondering is de dochter van onwetendheid.

Het is niet goed, ongelukkig te zijn, maar het is goed, het te zijn geweest.

Jonge vrouwen hebben niet genoeg humor en oudere vrouwen hebben niet genoeg schoonheid.

Om zonder tegenzin te worden gehoorzaamd, moet je het goede voorbeeld geven.

De angst voor de dood doet iemand alle kwaden en ongemakken van het leven vergeten.

De mooiste overwinning is jezelf te verslaan.

Kansrekening: De paradox van Chevalier de Méré

Antoine Gombaud, **Chevalier de Méré** (1607 – 1684), schrijver en verwoed gokker.

Spel 1: *Double or nothing* met een munt. Fifty-fifty (?)

Spel 2: Als je in vier worpen met een dobbelsteen een zes gooit wordt je inzet verdubbeld, anders ben je je geld kwijt.

Redenering van De Méré: Met één worp gooi je één op de zes keer een zes. Met vier dobbelstenen gooi je dus vier op de zes keer een zes. Dat is meer dan drie op de zes, dus hier kun je geld mee verdienen.

Spel 3: Als je in 24 worpen met twee dobbelstenen een keer twee zessen gooit wordt je inzet verdubbeld, anders ben je je geld kwijt.

Redenering van De Méré: Met één worp gooi je één op de 36 keer twee zessen. Met 24 worpen gooi je dus 24 op de 36 keer twee zessen. Dat is ook 4 op de 6, dus meer dan 1 op de 2, dus dit is ook een gunstig spel.

Kansrekening: De paradox van Chevalier de Méré

Spel 2: Als je in vier worpen met een dobbelsteen een zes gooit wordt je inzet verdubbeld, anders ben je je geld kwijt.

Fout in de redenering van De Méré: Met deze redenering zou je zés op de zes keer een zes gooien. Dat betekent dat je met zes worpen altijd een zes gooit!

Met twee worpen is de kans om een zes te gooien niet $2 \times \frac{1}{6} = \frac{2}{6} = 0{,}3333$ (De Méré)

maar $1 - \left(\frac{5}{6}\right)^2 = \mathbf{\frac{11}{36}} = 0{,}3056$ (Pascal)
(één min de kans dat je geen zes gooit)

$2^e \rightarrow$ $1^e \downarrow$	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

Kansrekening: De paradox van Chevalier de Méré

	Winstkans volgens De Méré	Winstkans volgens Pascal
Spel 1	$\frac{1}{2} = 0{,}5$	$\frac{1}{2} = 0{,}5$
Spel 2	$\frac{4}{6} = 0{,}6667$	$1 - \left(\frac{5}{6}\right)^4 = \frac{671}{1296} = 0{,}5177$
Spel 3	$\frac{24}{36} = 0{,}6667$	$1 - \left(\frac{35}{36}\right)^{24} = 0{,}4914$

Chevalier de Méré
(1607 – 1684)

Blaise Pascal
(1623 – 1662)

Spel 1 De speler speelt op termijn quitte.

Spel 2 De speler wint op termijn (gemiddeld 3,55% van de inzet).

Spel 3 verliest de speler op termijn (gemiddeld 1,72% van de inzet).

Bij roulette verlies je gemiddeld 2,78% van de inzet per spel

Kansrekening: Wat is een kans? (3.2.1)

Kansdefinitie van Pascal:

$$\text{Kans op een (samengestelde) gebeurtenis} = \frac{\text{aantal gunstige (elementaire) gebeurtenissen}}{\text{totaal aantal(elementaire)gebeurtenissen}}$$

Beperkingen:
- Het gaat om eindig veel gebeurtenissen
- Alle (elementaire) gebeurtenissen zijn even waarschijnlijk.

Discrete kansverdelingen - Begrippen

Een **kansexperiment** is een gebeurtenis (experiment, meting, afnemen van een enquête, …) die resulteert in een uitkomst (getal, gegevens) die niet van tevoren kan worden bepaald, maar onderhevig is aan een bepaalde mate van willekeur, bijvoorbeeld het aantal ogen gegooid met een dobbelsteen, een ingevuld enquêteformulier, kaliumgehalte in het bloed van een patiënt, of zijn/haar/het's gewicht.

Een **uitkomstenruimte** is de verzameling van alle mogelijke uitkomsten van een kansexperiment. In de discrete situatie is dat meestal een eindige, maar in ieder geval een aftelbare verzameling.

Een **kansvariabele** (= toevalsvariabele = stochastische variabele = **stochast**) is een eigenschap van de uitkomst van kansexperiment die in een getal is uit te drukken (bijvoorbeeld de lengte van een militair, het aantal slachtoffers van een IED, wachttijd voor een douanecontrole). Notatie: $\underline{k}$ of $\underline{x}$.
Als eindig of aftelbaar veel waarden worden aangenomen heet de stochast **discreet** ($\underline{k}$), anders **continu** ($\underline{x}$).

Een **kansfunctie** is een functie die aan elke waarde van de stochast een kans toekent, d.w.z. een getal tussen 0 en 1 (0 en 1 zelf mogen ook). $\boldsymbol{f(k) = P(\underline{k} = k)}$
De som van alle kansen in de kansruimte moet gelijk zijn aan 1.

De combinatie van kansexperiment, uitkomstenruimte en kansfunctie heet een **kansruimte**.

Discrete kansverdelingen - Begrippen

Voorbeeld: Het **kansexperiment** is een vraag uit een enquête naar het aantal kinderen van de ondervraagde.

De **uitkomstenruimte** is een beetje problematisch, want wat is het grootste aantal kinderen dat een man kan krijgen? Volgens het Guinness Book of World Records zou Moulay Ismaël Ibn Sharif, sultan van Marokko, de vader zijn van maar liefst 888 kinderen. Maar met kunstmatige inseminatie zou een zaaddonor in theorie miljoenen kinderen kunnen krijgen. Voor alle zekerheid kiezen we de verzameling van de natuurlijke getallen:

$$\mathbb{N} = \{1, 2, 3, 4, \ldots\ldots\}$$

Als **kansvariabele** nemen we het aantal kinderen $\underline{k}$ van de ondervraagde.
De **kansfunctie** zou kunnen zijn (fictief)

$$f(0) = P(\underline{k} = 0) = 0{,}158$$
$$f(1) = P(\underline{k} = 1) = 0{,}184$$
$$f(2) = P(\underline{k} = 2) = 0{,}436$$
$$f(3) = P(\underline{k} = 3) = 0{,}131$$
$$f(4) = P(\underline{k} = 4) = 0{,}067$$
$$f(5) = P(\underline{k} = 5) = 0{,}011$$
.....

Kansen zijn altijd getallen tussen 0 en 1.
De som van alle kansen is altijd 1

Een kansfunctie heeft dus altijd waarden tussen 0 en 1 en de som van alle waarden van de kansfunctie is altijd 1.

Uitleg van de notatie $f(k) = P(\underline{k} = k)$

1. Een =-teken wordt in de wiskunde op verschillende manieren gebruikt:
   a. Als toewijzing (*assignment*): Kies $x = 3$. In een formule met x wordt de waarde 3 ingevuld voor x.
   b. Als een vergelijking: De vergelijking "$2x - 1 = 3$" kun je herleiden tot (oplossen als) $x = 2$, d.w.z. linker- en rechterkant zijn alleen gelijk als $x = 2$.
   c. Als een vergelijkingsuitspraak: De uitspraak "$2x - 1 = 3$" zegt dat de linker- en rechterkant aan elkaar gelijk zijn. Die uitspraak is juist als $x = 2$ en onjuist als $x \neq 2$.

2. Verschil tussen $\underline{k}$ en k. Met $\underline{k}$ wordt een (discrete) kansvariabele bedoeld, bijvoorbeeld het aantal ogen dat met een dobbelsteen is gegooid. Na elk experiment (worp) levert $\underline{k}$ dus een getal op. $\underline{k}$ is niet een getal, maar stelt de uitkomst van een experiment voor, dat na elk experiment anders kan zijn. In plaats van $\underline{k}$, kan er bv. ook staan $\underline{dobbelsteenworp}$.
k is een variable die een getal voorstelt, hiervoor wordt dezelfde letter k gebruikt om aan te geven dat het bij de stochast $\underline{k}$ hoort, maar het had ook n of $aantal\_ogen$ kunnen heten. Je kunt bijvoorbeeld $k = 3$ kiezen, dan is $\underline{k} = 3$ de vraag: Levert het kansexperiment 3 ogen op? (het antwoord kan ja of nee zijn).

3. $\underline{k} = k$ moet je opvatten als een uitspraak (vergelijking) die waar is of niet, en $P(\underline{k} = 3)$ is de **kans** (Probability) dat de waarde die een worp oplevert (nl. $\underline{k}$) toevallig gelijk is aan 3. Voor een dobbelsteen geldt dus $P(\underline{k} = 3) = {}^1/_6$.

4. $f(k) = P(\underline{k} = k)$ betekent dat f een functie is (de kansfunctie), die aan waarden van k telkens een getal toevoegt, namelijk de kans dat het kansexperiment de waarde k oplevert. Deze kans wordt genoteerd als $P(\underline{k} = k)$

Kansfunctie Voorbeeld

Voorbeeld: Je gooit twee keer met een dobbelsteen. De uitkomst is de som van de ogen van de twee worpen. Dit is het kansexperiment.

Kansfunctie Voorbeeld

Voorbeeld: Je gooit twee keer met een dobbelsteen. De uitkomst is de som van de ogen van de twee worpen. Dit is het kansexperiment.
De kansvariabele $\underline{k}$ geeft het aantal ogen van de twee dobbelstenen samen.
De uitkomstenruimte bestaat uit alle paren $S = \{(n,m) \mid n, m = 1,2,3, \ldots, 6\} =$ $\{(1,1), (1,2), \ldots, (1,6), (2,1), (2,2), (2,3), \ldots, (2,6), (3,1), \ldots, (6,5), (6,6)\}$.
Er zijn 6 x 6 = 36 mogelijke uitkomsten.
De kansfunctie (welke kans hoort bij welke uitkomst?) kun je in een tabel geven →
of als een formule, bv.:

$$f(k) = \frac{6-|k-7|}{36}$$

k	$f(k)$
2	$^1/_{36}$
3	$^2/_{36}$
4	$^3/_{36}$
5	$^4/_{36}$
6	$^5/_{36}$
7	$^6/_{36}$
8	$^5/_{36}$
9	$^4/_{36}$
10	$^3/_{36}$
11	$^2/_{36}$
12	$^1/_{36}$

2ᵉ→ 1ᵉ↓	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Tabel met alle mogelijke worpen (waarden van $\underline{k}$)

Intermezzo: Het begrip aftelbaarheid (Extra inzicht, niet verplicht)

Een verzameling heet **aftelbaar** als je de elementen van de verzameling kunt nummeren: 1, 2, 3, 4, … en ze zo allemaal krijgt.

Voorbeeld: Elke eindige verzameling is aftelbaar.

Voorbeeld: De verzameling natuurlijke getallen $\mathbb{N} = \{0, 1, 2, 3, 4, 5, \dots\}$ is aftelbaar, maar ook oneindig groot.

Opgave: Zijn de gehele getallen $\mathbb{Z} = \{\dots, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, \dots\}$ aftelbaar?

Opgave: Zijn de rationale getallen (breuken) $\mathbb{Q} = \left\{\frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0\right\}$ aftelbaar?

Voorbeeld: De reële getallen $\mathbb{R}$ (alle kommagetallen) is niet aftelbaar (overaftelbaar).

Verwachtingswaarde, variantie, standaarddeviatie van een discrete kansverdeling

$\underline{k}$ is een kansvariabele met kansfunctie $f(k) = P(\underline{k} = k)$

Bij een kansvariabele kun je de volgende getallen uitrekenen:

De **verwachtingswaarde** van $\underline{k}$ is $\mu = E(\underline{k}) = \sum_{\text{alle } k} kf(k)$

De **variantie** van $\underline{k}$ is $\text{Var}(\underline{k}) = E\left((\underline{k} - \mu)^2\right) = \sum_{\text{alle } k} (k - \mu)^2 f(k)$

Rekenregel: $\text{Var}(\underline{k}) = E(\underline{k}^2) - \mu^2 = \sum_{\text{alle } k} k^2 f(k) - \mu^2$

De **standaarddeviatie** van $\underline{k}$ is $\sigma(\underline{k}) = \sqrt{\text{Var}(\underline{k})}$

k	f(k)	kf(k)	k^2 f(k)
2	$^1/_{36}$		
3	$^2/_{36}$		
4	$^3/_{36}$		
5	$^4/_{36}$		
6	$^5/_{36}$		
7	$^6/_{36}$		
8	$^5/_{36}$		
9	$^4/_{36}$		
10	$^3/_{36}$		
11	$^2/_{36}$		
12	$^1/_{36}$		
som	1		

Verwachtingswaarde, variantie, standaarddeviatie

$\underline{k}$ is een kansvariabele met kansfunctie $f(k) = P(\underline{k} = k)$

De **verwachtingswaarde** van $\underline{k}$ is $\mu = E(\underline{k}) = \sum_{\text{alle } k} kf(k)$

De **variantie** van $\underline{k}$ is $\mathrm{Var}(\underline{k}) = E\left((\underline{k} - \mu)^2\right) = \sum_{\text{alle } k} (k - \mu)^2 f(k)$

Rekenregel: $\mathrm{Var}(\underline{k}) = E(\underline{k}^2) - \mu^2 = \sum_{\text{alle } k} k^2 f(k) - \mu^2$

De **standaarddeviatie** van $\underline{k}$ is $\sigma(\underline{k}) = \sqrt{\mathrm{Var}(\underline{k})}$

Handmatig uitrekenen met een tabel

k	f(k)	kf(k)	k^2 f(k)
2	$1/36$	$2/36$	$4/36$
3	$2/36$	$6/36$	$18/36$
4	$3/36$	$12/36$	$48/36$
5	$4/36$	$20/36$	$100/36$
6	$5/36$	$30/36$	$180/36$
7	$6/36$	$42/36$	$294/36$
8	$5/36$	$40/36$	$320/36$
9	$4/36$	$36/36$	$324/36$
10	$3/36$	$30/36$	$300/36$
11	$2/36$	$22/36$	$242/36$
12	$1/36$	$12/36$	$144/36$
som	1	7	$1974/36$

Verwachtingswaarde, variantie, standaarddeviatie

$\underline{k}$ is een kansvariabele met kansfunctie $f(k) = P(\underline{k} = k)$

De **verwachtingswaarde** van $\underline{k}$ is $\mu = E(\underline{k}) = \sum_{\text{alle } k} kf(k)$

De **variantie** van $\underline{k}$ is $\mathrm{Var}(\underline{k}) = E\left(\left(\underline{k} - \mu\right)^2\right) = \sum_{\text{alle } k}(k - \mu)^2 f(k)$

Rekenregel: $\mathrm{Var}(\underline{k}) = E(\underline{k}^2) - \mu^2 = \sum_{\text{alle } k} k^2 f(k) - \mu^2$

De **standaarddeviatie** van $\underline{k}$ is $\sigma(\underline{k}) = \sqrt{\mathrm{Var}(\underline{k})}$

Voorbeeld: Totaal aantal ogen van 2 dobbelstenen:

$$\mu = 7$$

$$\mathrm{Var}(\underline{k}) = {}^{1974}/_{36} - 7^2 = 5{,}8333\ldots$$

$$\sigma(\underline{k}) = \sqrt{5{,}8333} = 2{,}4152\ldots$$

k	f(k)	kf(k)	k^2 f(k)
2	$^1/_{36}$	$^2/_{36}$	$^4/_{36}$
3	$^2/_{36}$	$^6/_{36}$	$^{18}/_{36}$
4	$^3/_{36}$	$^{12}/_{36}$	$^{48}/_{36}$
5	$^4/_{36}$	$^{20}/_{36}$	$^{100}/_{36}$
6	$^5/_{36}$	$^{30}/_{36}$	$^{180}/_{36}$
7	$^6/_{36}$	$^{42}/_{36}$	$^{294}/_{36}$
8	$^5/_{36}$	$^{40}/_{36}$	$^{320}/_{36}$
9	$^4/_{36}$	$^{36}/_{36}$	$^{324}/_{36}$
10	$^3/_{36}$	$^{30}/_{36}$	$^{300}/_{36}$
11	$^2/_{36}$	$^{22}/_{36}$	$^{242}/_{36}$
12	$^1/_{36}$	$^{12}/_{36}$	$^{144}/_{36}$
som	1	7	$^{1974}/_{36}$

Operaties met kansvariabelen H4.4, 4.5

Operatie	$\underline{k}$	$\mu(\underline{k})$	Var$(\underline{k})$	$\sigma(\underline{k})$
Vermenigvuldigen met vast getal α	$\alpha\underline{k}$	$\alpha\mu(\underline{k})$	$\alpha^2\text{Var}(\underline{k})$	$\lvert\alpha\rvert\sigma(\underline{k})$
Vast getal β erbij tellen	$\underline{k}+\beta$	$\mu(\underline{k})+\beta$	Var$(\underline{k})$	$\sigma(\underline{k})$
Twee onafhankelijke stochasten optellen	$\underline{k}+\underline{l}$ als $\underline{k},\underline{l}$ onafhankelijk	$\mu(\underline{k})+\mu(\underline{l})$	Var$(\underline{k})$+Var$(\underline{l})$	$\sqrt{\sigma(\underline{k})^2+\sigma(\underline{l})^2}$
Twee onafhankelijke stochasten aftrekken	$\underline{k}-\underline{l}$ als $\underline{k},\underline{l}$ onafhankelijk	$\mu(\underline{k})-\mu(\underline{l})$	Var$(\underline{k})$+Var$(\underline{l})$	$\sqrt{\sigma(\underline{k})^2+\sigma(\underline{l})^2}$
Som van resultaten van herhaald experiment (zelfde stochast)	$\underline{k}_1+\underline{k}_2+\cdots+\underline{k}_n$	$n\cdot\mu(\underline{k})$	$n\cdot\text{Var}(\underline{k})$	$\sqrt{n}\sigma(\underline{k})$
Gemiddelde van resultaten van herhaald experiment	$\dfrac{\underline{k}_1+\underline{k}_2+\cdots+\underline{k}_n}{n}$	$\mu(\underline{k})$	$\frac{1}{n}\text{Var}(\underline{k})$	$\dfrac{1}{\sqrt{n}}\sigma(\underline{k})$

Twee stochasten $\underline{k}$ en $\underline{l}$ zijn onafhankelijk als $\text{P}(\underline{k}=k \text{ én } \underline{l}=l)=\text{P}(\underline{k}=k)\cdot\text{P}(\underline{l}=l)$ voor alle mogelijke k en l.

Let op: $\underline{k}_1+\underline{k}_2+\cdots+\underline{k}_n$ is niet hetzelfde als $n\underline{k}$, want een stochast $\underline{k}$ kan bij elke aanroep een andere waarde opleveren.

Let op: **Veelgemaakte fouten: $\boldsymbol{\sigma(\underline{k})+\sigma(\underline{l})}$ i.p.v. $\boldsymbol{\sqrt{\sigma(\underline{k})^2+\sigma(\underline{l})^2}}$ en $\boldsymbol{n\sigma(\underline{k})}$ of $\boldsymbol{\sqrt{n\sigma(\underline{k})}}$ i.p.v. $\boldsymbol{\sqrt{n}\sigma(\underline{k})}$.**

Operaties met kansvariabelen

dag	μ	σ
Ma	22,2	3,1
Di	18,7	2,2
Wo	21,3	3,0
Do	16,7	2,7
Vr	24,0	1,9
Za	33,1	2,9
Zo	35,7	4,2

Voorbeeld: Een hotel heeft 40 kamers ter beschikking. Het gemiddeld aantal kamers dat per dag wordt verhuurd (μ) en de standaarddeviatie (σ) daarin zijn in het verleden bepaald en staan in de tabel hiernaast.

a. De prijs van een kamer is 120 euro. De vaste lasten (personeel, huur, voorraden, voorzieningen, e.d.) op maandagen zijn 2100 euro. Bereken de gemiddelde winst op maandag, met de standaarddeviatie daarin.

Als $\underline{x_1}$ de kansvariabele is die het aantal verhuurde kamers op maandag voorstelt, dan is de kansfunctie (de kansen dat er op maandag 0, 1, 2, …, 40 kamers worden verhuurd) niet bekend, maar wel μ en σ. **Gebruik de rekenregels!**

De inkomsten op maandag zijn $120\underline{x_1} - 2100$ euro.

De gemiddelde waarde daarvan is $120 \cdot 22{,}2 - 2100 = 564$ euro.

De standaarddeviatie is dan $120 \cdot 3{,}1$ = 372 euro.

Operaties met kansvariabelen

dag	μ	σ
Ma	22,2	3,1
Di	18,7	2,2
Wo	21,3	3,0
Do	16,7	2,7
Vr	24,0	1,9
Za	33,1	2,9
Zo	35,7	4,2

Voorbeeld: Een hotel heeft 40 kamers ter beschikking. Het gemiddeld aantal kamers dat per dag wordt verhuurd (μ) en de standaarddeviatie (σ) daarin zijn in het verleden bepaald en staan in de tabel hiernaast.

b. Bereken over vier achtereenvolgende maandagen, hoeveel kamers er naar verwachting in totaal zijn bezet, en de standaarddeviatie daarin.

Als $\underline{x_1}$ de kansvariabele is die het aantal verhuurde kamers op maandag voorstelt, dan is het aantal kamers dat op 4 dagen is verhuurd $\underline{x_1} + \underline{x_1} + \underline{x_1} + \underline{x_1}$.

Dat is wat anders dan $4\underline{x_1}$, want op elk van de vier maandagen kan de werkelijke waarde van $\underline{x_1}$ anders zijn, want het is een kansvariabele.

De verwachtingswaarde (rekenregel): $4\mu = 4 \cdot 22{,}2 = 88{,}8$ (kamers per nacht).

De standaarddeviatie daarvan is (rekenregel): $\sqrt{4}\sigma = 2 \cdot 3{,}1 = 6{,}2$.

Operaties met kansvariabelen

dag	μ	σ
Ma	22,2	3,1
Di	18,7	2,2
Wo	21,3	3,0
Do	16,7	2,7
Vr	24,0	1,9
Za	33,1	2,9
Zo	35,7	4,2

Voorbeeld: Een hotel heeft 40 kamers ter beschikking. Het gemiddeld aantal kamers dat per dag wordt verhuurd (μ) en de standaarddeviatie (σ) daarin zijn in het verleden bepaald en staan in de tabel hiernaast.

c. Bereken de gemiddelde bezettingsgraad (als percentage) in het weekend, en de standaarddeviatie daarin.

Als $\underline{x_6}$ de kansvariabele is die het aantal verhuurde kamers op zaterdag voorstelt, en $\underline{x_7}$ die op zondag, dan is de bezettingsgraad in het weekend $100\frac{\underline{x_6}+\underline{x_7}}{40+40} = 1{,}25\left(\underline{x_6}+\underline{x_7}\right)$ procent.

De verwachtingswaarde daarvan is $1{,}25(\mu_6+\mu_7) = 1{,}25(33{,}1+35{,}7) = 86\%$.

De standaarddeviatie daarin is $1{,}25\sqrt{\sigma_6^2+\sigma_7^2} = 1{,}25\sqrt{2{,}9^2+4{,}2^2} = 6{,}38\%$.

Operaties met kansvariabelen

dag	μ	σ
Ma	22,2	3,1
Di	18,7	2,2
Wo	21,3	3,0
Do	16,7	2,7
Vr	24,0	1,9
Za	33,1	2,9
Zo	35,7	4,2

Voorbeeld: Een hotel heeft 40 kamers ter beschikking. Het gemiddeld aantal kamers dat per dag wordt verhuurd (μ) en de standaarddeviatie (σ) daarin zijn in het verleden bepaald en staan in de tabel hiernaast.

d. In het weekend worden kamers voor één nacht verhuurd, of voor het weekend. Wat kun je zeggen over het gemiddeld aantal kamers dat het hele weekend aan één klant wordt verhuurd?

Dat zit tussen 0 (alle kamers worden los verhuurd) en 33,1 (op zaterdag worden geen kamers los verhuurd en op zondag gemiddeld 35,7 – 33,1 = 2,6).

Bernoulliverdeling (“Alternatieve verdeling”) H6.2

De eenvoudigste discrete kansverdeling gaat over een kansexperiment met maar twee mogelijke uitkomsten: 0 en 1 (kop/munt, ja/nee), die optreden met een vaste kans, bv.: Een punaise opgooien.

Jakob Bernoulli
(1654 – 1705)

Dit wordt beschreven door een kansvariabele $\underline{k}$ met mogelijke waarden 1 (slagen) en 0 (niet slagen) en:

$$P(\underline{k}=1)=p \quad \text{en} \quad P(\underline{k}=0)=1-p$$

$p \in [0,1]$, de slaagkans, is een willekeurig, vast getal
(Bv. $p=\frac{1}{2}$ voor een eerlijke munt, of $p=\frac{1}{790.499}$, kans op minstens een miljoen bij de BankGiro loterij) BankGiro Loterij biedt meeste kans op miljoenenprijs | Consumentenbond

De verwachtingswaarde van de Bernoulliverdeling is

$$\mu(\underline{k})=E(\underline{k})=\sum_{k=0}^{1} k\cdot P(\underline{k}=k)=0\cdot P(\underline{k}=0)+1\cdot P(\underline{k}=1)=p$$

De variantie is $\sigma^2=\mathrm{Var}(\underline{k})=E(\underline{k}^2)-\mu^2$

$$=\left(0^2\cdot P(\underline{k}=0)+1^2\cdot P(\underline{k}=1)\right)-p^2=p(1-p)$$

Standaardeviatie: $\sigma=\sqrt{p(1-p)}$

Bernoulliverdeling

k	f(k)	kf(k)	k^2 f(k)
0	$1-p$	0	0
1	p	p	p
som	1	p	p

$$\mu=p$$

$$\sigma=\sqrt{p(1-p)}$$

Er is één parameter: p

Binomiale verdeling H6.1 + 6.2

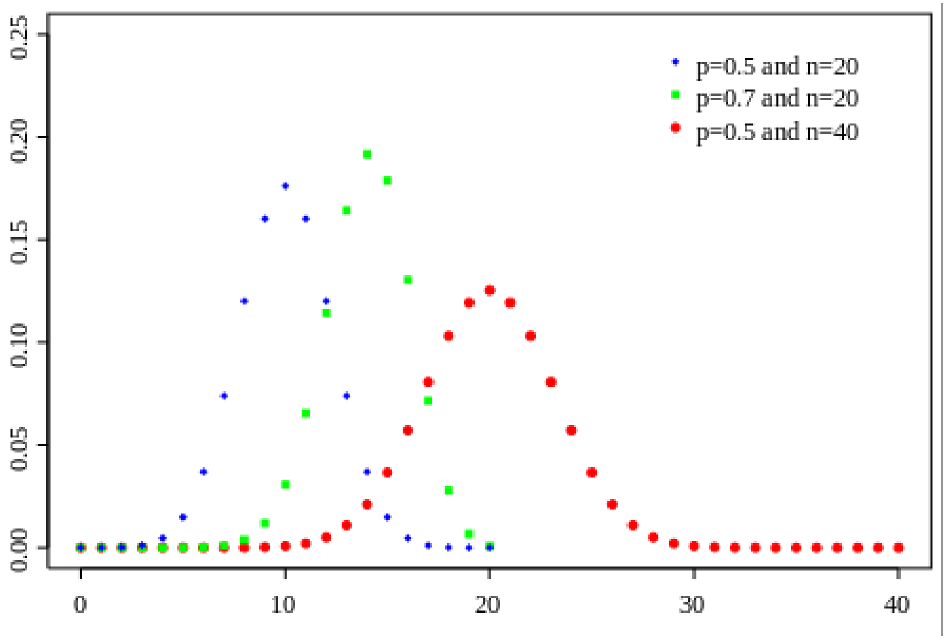
De **binomiale verdeling** beschrijft een kansexperiment met net als bij de Bernoulliverdeling twee mogelijke uitkomsten, (0 en 1) maar je voert dit vervolgens n keer uit, steeds met dezelfde slaagkans p. De kansvariabele $\underline{k}$ is het aantal keren met succes (uitkomst 1 met kans p). Er geldt

$$\underline{k} = \underline{k}_1 + \underline{k}_2 + \cdots + \underline{k}_n,$$

waarbij $\underline{k}_j$ steeds het antwoord 0 of 1 levert volgens de Bernoulliverdeling. $\underline{k}$ wordt beschreven door de kansfunctie:

$$f(k) = P(\underline{k} = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{voor} \quad k = 0, 1, 2, \dots, n$$

$p \in [0,1]$ is een willekeurig vast getal (Bv. $p = \frac{1}{2}$ voor een eerlijke munt).



Omdat de trekkingen onafhankelijk zijn is verwachtingswaarde de som van de verwachtingswaarden van n keer het Bernoulli-experiment (zie rekenregels):

$$\mu = E(\underline{k}) = nE(\underline{k}_1) = np$$

Hetzelfde geldt voor de variantie:

$$\sigma^2 = E\left((\underline{k} - \mu)^2\right) = np(1-p), \text{ dus}$$

$$\sigma = \sqrt{np(1-p)}.$$

Binomiale verdeling

$$\mu = np$$

$$\sigma = \sqrt{np(1-p)}$$

Twee parameters: n en p

Notatie: $\underline{k} \sim \text{bin}(n; p)$

Binomiale verdeling: Afleiding m.b.v. differentiëren (Extra info, niet voor tentamen)

Er geldt (Binomium van Newton):

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Kies $a = p, \quad b = 1 - p$ levert:

$$1 = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

Differentieer het Binomium naar a en vermenigvuldig dan links en rechts met a:

$$na(a+b)^{n-1} = \sum_{k=0}^{n} k \binom{n}{k} a^k b^{n-k}$$

Vul weer $a = p, \quad b = 1 - p$ in:

$$E(\underline{k}) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np$$

Nog een keer dezelfde truc levert:

$$E(\underline{k}^2) = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} = np + n(n-1)p^2 = \boldsymbol{n^2 p^2}$$

dus

$$\mathrm{Var}(\underline{k}) = E(\underline{k}^2) - E(\underline{k})^2 = np + n(n-1)p^2 - n^2p^2 = np(1-p)$$

Voorbeeld

Neem aan dat $\underline{k} \sim \text{bin}(n = 15; p = 0{,}4)$, dan kun je $P\left(\underline{k} = 10\right)$ als volgt uitrekenen:

1. Grafische rekenmachine (DISTR): binompdf(15, 0.4, 10) = 0,0245
2. Grafische rekenmachine (MATH PRB nCr): $\binom{15}{10} \cdot 0{,}4^{10} \cdot (1 - 0{,}4)^{15-10}$= 0,0245
3. Tabel C3 boek: $P\left(\underline{k} = 10\right) = P\left(\underline{k} \leq 10\right) - P\left(\underline{k} \leq 9\right) = 0{,}0245$ Doe dit niet!!!
4. Excel: = BINOM.DIST(10;15;0.4;FALSE) = 0,024486
   = BINOM.VERD(10;15;0,4;ONWAAR) = 0,024486

Binomiale verdeling – binompdf en binomcdf

Op de grafische rekenmachine (TI84) staan onder DISTR twee binomiale functies:

binom**pdf** (n, p, x) (= Probability Density Function = kansfunctie) en
binom**cdf** (n, p, x) (= Cumulative Density Function = verdelingsfunctie)

Als $\underline{k} \sim \text{bin}(n = 15; p = 0{,}4)$ dan is

$P(\underline{k} = 10)$ = binom**p**df(15, 0.4, 10) = 0,0245

$P(\underline{k} \leq 10)$ = binom**c**df(15, 0.4, 10) = 0,9907

$P(\underline{k} \geq 10)$ = 1 – $P(\underline{k} < 10) = 1 - P(\underline{k} \leq 9) = 1 -$binom**c**df(15, 0.4, 9) = 0,0338

$P(6 \leq \underline{k} \leq 10)$ = $P(\underline{k} \leq 10) - P(\underline{k} \leq 5) =$
= binom**c**df(15, 0.4, 10) – binom**c**df(15, 0.4, 5) = 0,9907 – 0,4032 = 0,5874

Binomiale verdeling. Voorbeeld: Operation Eagle Claw

In 1979 werd de US ambassade in Teheran bestormd door volgelingen van ayatollah Khomeini en meer dan een jaar lang bezet gehouden. In de ambassade bevond zich 52 US gegijzelden.
Voorjaar van 1980 werd een bevrijdingsactie ondernomen (Operation Eagle Claw). Voor uitvoering van het tactische plan waren minimaal 6 helikopters nodig die vanuit vliegdekschip USS Nimitz werden ingezet.

De kans op volledig operationele inzetbaarheid gedurende de hele missie werd voor elke helikopter ingeschat op 0,8. Men besloot voor de zekerheid 8 helikopters in te zetten. Hoe groot was de theoretische kans op een succesvolle afloop?

De operatie mislukte. Twee Sea Stallion helikopters kwamen in problemen door een zandstorm, een derde had problemen met het hydraulische systeem (.

Hoeveel helikopters had men moeten inzetten om een succeskans van 0,99 voor de operatie te realiseren?

Binomiale verdeling. Operation Eagle Claw

- 6 operationele helikopters zijn vereist
- Kans operationeel is 0,8 per helikopter

1. Wat is de succeskans?

2. Hoeveel helikopters nodig voor succeskans 0,99?

1. Kies het aantal operationele helikopters $\underline{k} \sim \text{bin}(n = 8; p = 0{,}8)$, dan is de succeskans

$$P(\underline{k} \geq 6) = 1 - P(\underline{k} \leq 5) = 1 - \text{binomcdf}(8, 0.8, 5) = \mathbf{0,7969}$$

2. Veronderstel dat $\underline{k} \sim \text{bin}(n = ?\,; p = 0{,}8)$ en zoek n zodanig dat

$$P(\underline{k} \geq 6) = 1 - P(\underline{k} \leq 5) \geq 0{,}99, \qquad \text{ofwel} \qquad P(\underline{k} \leq 5) \leq 0{,}01$$

Trail and error (maak een tabel met de GR):

$$\text{binomcdf}(11, 0.8{,}5) = 0{,}01165$$
$$\text{binomcdf}(12, 0.8{,}5) = 0{,}00390$$

Dus $n = 12$ helikopters nemen.