Oefening 1 (**) - Bewijs dat de absolute waarde van de correlatiecoëfficiënt ρ niet groter kan zijn dan 1. Hint: Vertrek van de variantie van $x-y$, of gebruik de ongelijkheid van Cauchy-Schwartz: $(\sum_i u_i v_i)^2 \leq (\sum_i u_i)^2 (\sum_i v_i)^2$ voor twee vectoren $\vec{u}$ en $\vec{i}$.

Oefening 2 (*) - Als we het n-de moment van een verdeling voorstellen door $\alpha_n = \overline{x^n}$, en μ_n is het n-de centrale moment, en $m = \bar{x}$ is het gemiddelde, bewijs dat :

$$\mu_3 = \alpha_3 - 3m\alpha_2 + 2m^3$$

Oefening 3 (*) - Simpson's paradox op een loterij wordt 240 euro verdeeld in 10 enveloppen. Je kan een blauwe enveloppe kiezen voor 10 euro . Een rode enveloppe kost 20 euro. In de 5 blauwe enveloppen vinden we respectievelijk 5,5,10,10 of 50 euro. In de 5 rode enveloppen zit 10,10,20,20 of 100 euro.

1. Wat is de gemiddelde winst als je een blauwe enveloppe kiest? En voor de rode ?

2. Nu werd beslist om één van de 20 euro biljetten die in een rode enveloppe zaten naar een extra blauwe enveloppe te verplaatsen. Welke groep is door die verruiling bevoordeeld?

3. Hoe is dat te verklaren?

Oefening 4 (**) - Een stralingsbron bevat een aantal instabiele kernen. Elke seconde vervallen er een aantal met uitzending van straling. Het aantal kernen dat per seconde vervalt noemen we de activiteit van de bron. De eenheid daarvan is de becquerel (Bq) : 1Bq = 1 vervalsproces per seconde. Een experiment tracht de activiteit van een bron te bepalen. Gedurende een week zijn 12 metingen van vervallen uitgevoerd.

tijd (dag)	0.97	1.39	1.69	2.91	3.24	3.95	4.44	4.47	4.56	5.68	5.75	5.99
activiteit (GBq)	31184	13667	7414	654	337	81	31	29	24	3	2	1

1. Bereken de correlatiecoëfficiënt tussen het moment van de meting en de gemeten activiteit. Verklaar dit resultaat.

2. Teken het spreidingsdiagram op een slimme manier. Is het mogelijk om een betere correlatiecoëfficiënt uit te werken?

Oefening 5 (*) - Bewijs de gelijkwaardigheid van de drie uidrukkingen in de definitie van de covariantie (Definitie 1.3.1):

$$\mathrm{cov}(x,y) = \overline{(x-\bar{x})(y-\bar{y})} = \overline{xy} - \bar{x}\bar{y}$$

Oefening 6 (*) - De temperatuur (in graden Celcius) in een experimentele installatie met een modulaire structuur werd gemeten op de acht verschillende modulen met als resultaten:

module	1	2	3	4	5	6	7	8
Temp.(C)	40	37	43	46	39	43	35	37

1. Bereken de steekproefvariantie s^2.

2. Indien de temperatuur uit de steekproef in graden Kelvin was gemeten, wat zou de variantie zijn?

3. Indien de temperatuur in graden Fahrenheit was gemeten, waaraan is s^2 dan gelijk? ($C = \frac{5}{9}(F - 32)$)

4. Is er een correlatie tussen de positie van de module en de gemeten temperatuur?

Statistiek: Oefeningenles 2

Oefening 1 (*) - Mongoolse moeraskoorts is een ziekte (Z) die weinig voorkomt. Een arts maakt ze maar mee in 1 op 10 000 patiënten. Iemand die er aan lijdt vertoont altijd vlekjes (V) en acute lethargie (L). Over het algemeen (in 60% van de gevallen) is er ook een onlesbare dorst (D), en 1 op de 5 patiënten heeft hevige niesbuien (N). Deze symptomen kunnen echter ook het gevolg zijn van andere ziekten: zo vindt men dat bij mensen die niet aan Mongoolse moeraskoorts lijden 3% vlekjes heeft, 10% lethargisch is, 2% dorst heeft en 5% zwaar niest. De waarschijnlijkheid voor het optreden van deze symptomen is onafhankelijk.

1. Toon aan dat als je naar de dokter gaat met al deze symptomen, de kans dat je ook werkelijk lijdt aan Mongoolse moeraskoorts 80 % bedraagt.

2. Wat is de kans als je ze allemaal hebt behalve niezen?

Oefening 2 (*/) -** Stel je een kansexperiment voor dat bestaat uit het opgooien van een rode en een blauwe dobbelsteen. De 6 zijden van de rode dobbelsteen zijn $\{1, 1, 1, 2, 2, 3\}$. Die van de blauwe zijn $\{1, 2, 2, 3, 3, 4\}$.

1. Wat is de waarschijnlijkheidsverdeling van de toevallige variabele R (en B), gedefinieerd voor de mogelijke uitkomsten van de rode (en blauwe) dobbelsteen? Wat zijn de respectievelijke verwachtinsgwaarden?

2. Wat is de waarschijnlijkheidsverdeling van de toevallige variabele S, de som van de twee dobbelstenen? Geef zijn verwachtingswaarde.

3. Wat is de gezamenlijke waarschijnlijkheidsverdeling van R en B?

4. Stel de variabele M als de maximale waarde van de twee dobbelstenen. Wat is zijn waarschijnlijkheidsverdeling?

5. Nu gooien we eerst de rode dobbelsteen. Zijn waarde R definieert het aantal worpen met de blauwe dobbelsteen. X is gedefinieerd als de som van de R uitkomsten van de blauwe dobbelsteen. Geef de waarschijnlijkheidsverdeling van X. Wat is de correlatiecoefficiënt tussen X en R?

Oefening 3 (**) - Stel dat voor een toevallige variabele X en een zekere $\alpha > 0$

$$P(X > x) = x^{-\alpha}$$

voor $x > 1$. Bereken de dichtheid van X, zijn verwachtingswaarde en variantie (indien deze getallen bestaan).

Oefening 4 (*) - Stel dat de dichtheid van de discrete toevallige variabelen (X, Y) gegeven is door:

$$P(X = k_1; Y = k_2) = \begin{cases} 1/3 & (k_1, k_2) \in [(0,0), (1,1), (1,-1)] \\ 0 & \text{elders} \end{cases}$$

Bewijs dat $\rho_{XY} = 0$ maar dat X en Y niet onderling onafhankelijk zijn.

Oefening 5 (**) - Gegeven de gezamenlijke dichtheid van twee continue toevallige variabelen $X \in [0,2]$ en $Y \in [0,1]$:

$$f_{XY}(x;y) = \begin{cases} kxy & 2y < x \\ 0 & \text{elders} \end{cases}$$

1. Bepaal de waarde van k.
2. Bereken $P(X < 1, Y < 0.5)$
3. Bepaal de marginale en conditionele waarschijnlijkheidsverdelingen van X en Y.
4. Bereken $E(Y|x)$
5. Zijn X en Y onafhankelijk?

Statistiek: Oefeningenles 3

Oefening 1 (*) Bewijs (stelling 3.1.1 in de cursus) dat de exponenentiële dichtheid geheugenloos is:

$$P(X > x_1 + x_2 | X > x_2) = P(X > x_1)$$

Oefening 2 (*/)** Stel dat X $B(p = 0.5, n = 16)$ (d.w.z. verdeeld volgens een binomiaalverdeling, met p de kans op succes, en n het aantal pogingen). Bereken $P(X \leq 5)$ op 4 verschillende manieren.

1. exact
2. via de Poissonbenadering
3. via de Gaussische verdeling zonder continuïteitscorrectie
4. via de Gaussische verdeling met continuïteitscorrectie

Welke benadering werkt in dit geval het best?

Oefening 3 (*) Als maat voor de breedte van een verdeling kan je, naast de standaardafwijking σ, ook de FWHM gebruiken. Bewijs (stelling 1.2.4 in de cursus) dat voor de Gaussverdeling:

$$\text{FWHM} = 2\sigma\sqrt{2\ln 2} = 2.35\sigma$$

Oefening 4 (**) Een lifter staat langs een weg waar gemiddeld 1 auto per minuut passeert. De waarschijnlijkheid dat een chauffeur hem een lift geeft is 1%. Wat is de kans dat de lifter er nog altijd staat

1. nadat er 60 wagens gepasseerd zijn?
2. na 1 uur?

Oefening 5 (**) De Gammaverdeling heeft twee positieve parameters: de vormparameter k en de schaalparameter θ. Haar waarschijnlijkheidsdichtheid is gedefinieerd voor $x \geq 0$, als

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

waar Γ de gammafunctie voorstelt:

$$\Gamma(x) = \int_0^{+\infty} e^{-t} t^{x-1} \mathrm{d}t$$

1. De gammafunctie is een uitbreiding van de welbekende faculteitfunctie voor de natuurlijke getallen. Bewijs dat als x een posititief, reëel getal is, $\Gamma(x) = (x-1)\Gamma(x-1)$.
2. Bepaal de verwachtingswaarde en de variantie van de Gammadichtheid f.
3. De exponentiële verdeling is een speciaal geval van de Gammaverdeling. Zoek de parameters waarvoor deze twee verdelingen gelijk zijn, en ga na dat de eerste twee momenten (hierboven berekend) overeenkomen.

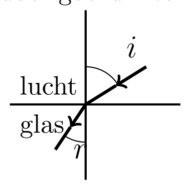
Oefening 6 (**) Stel dat de weerstand van een set resistoren van hetzelfde model normaal verdeeld is. 10% van deze resistoren hebben een weerstand die groter is dan 10.256 Ω, en 5% hebben een weerstand kleiner dan 9.671 Ω. Wat zijn het gemiddelde en de standaardafwijking van deze verdeling?

Oefening 7 (***) Bewijs de formules voor de transformatie tussen een gecorreleerde $(\sigma_x, \sigma_y, \rho)$ en een ongecorreleerde parametrisatie (σ_u, σ_v) van een binormale verdeling, indien de rotatiehoek van de transformatiematrix θ bedraagt. Zie einde hoofdstuk 3.

Statistiek: Oefeningenles 4

Oefening 1 (*) Je meet de brekingsindex van glas door gebruik te maken van de wet van Snellius:

$$\sin i = n \sin r$$



en je vindt de volgende waarden:

i	r
$(20\pm1)^\circ$	$(13\pm1)^\circ$
$(40\pm1)^\circ$	$(23.5\pm1)^\circ$

1. Geef de brekingsindex $n \pm \sigma_n$.
2. Zijn de meetresultaten compatibel?

Oefening 2 (*) Een radioactieve bron (met lange levensduur) geeft 389 tellen in de eerste minuut en 423 in de tweede. Wat is het beste gecombineerde resultaat?

Oefening 3 (/***)** We willen de omtrek L en de oppervlakte A (met respectievelijke fouten) van een willekeurige driehoek berekenen die gekarakteriseerd wordt door een hoek α en de lengte van zijn twee aangrenzende zijden a en b. De fouten op a, b en α zijn ongecorreleerd.

1. Bepaal de fouten op L en A en de correlatie ertussen.
2. Pas de gevonden formules toe op deze meting:

α	$62^\circ 12' 32'' \pm 7''$
a	100.23 ± 0.05 m
b	87.12 ± 0.04 m

3. Stel dat de meting van de lengte een systematische fout S van 0.5cm heeft. Wat zijn de gevolgen daarvan in de covariantiematrix?

Statistiek: Oefeningenles 5

Oefening 1 (*) Een binomiaal experiment geeft je N_S successen en N_F mislukkingen. Bewijs dat

$$\hat{p} = \frac{N_S}{N_S + N_F}$$

een consistente, onbevooroordeelde schatting is van de individuele waarschijnlijkheid p.

Oefening 2 (**) Veronderstel dat $X_1, ..., X_n$ een steekproef is genomen uit een Poisson-verdeling. Men wenst λ^2 te schatten.

1. Bewijs dat $(\bar{X})^2$ geen onbevooroordeelde schatter voor λ^2 is.
2. Vind een schatter die wel onbevooroordeeld is

Oefening 3 (**) Gegeven de zeven getallen

20.0 19.7 20.6 18.5 21.2 20.8 20.7

1. Bereken het gemiddelde. Wat is de fout op het gemiddelde als de getallen verschillende metingen voorstellen van één (Gauss) verdeelde grootheid, met een resolutie van 0.8?
2. Schat de standaardafwijking als je weet dat het echte gemiddelde 20.0 is. Wat is de fout op de schatting ?
3. Schat de standaardafwijking zonder voorafgaande kennis van het echte gemiddelde. Wat is de fout op de schatting ?
4. Bepaal de fout op het gemiddelde als je geen voorafgaande kennis hebt van de standaardafwijking.

Oefening 4 (/***)** **Achtergrond:** Het aantal manieren om een steekproef van n gebeurtenissen (met gelijke kansen) te selecteren in een set van N mogelijkheden is gegeven door C_n^N. Elke gebeurtenis wordt in verband gebracht met een gegeven variabele X. Over de hele set (N mogelijkheden) hebben we dus een gemiddelde waarde $\bar{X}$, die voor een bepaalde steekproef (n gebeurtenissen) via een schatter $\hat{\mu}$ van $\bar{X}$ kan geschat worden. Die schatter is onbevooroordeeld. De variantie van de schatter is gegeven door :

$$V[\hat{\mu}] = \frac{N-n}{N-1}\frac{\sigma_X^2}{n}$$

Vraagstuk: Een arts wil het gemiddelde aantal cigaretten m bepalen, dat per dag door een student gerookt wordt. Daarvoor stuurt hij een enquete rond, en slechts 30 studenten hebbenn gereageerd. De resulaten zijn beschikbaar per studiejaar (veronderstel dat de σ voor elke groep studenten gekend is).

Studiejaar	aantal studenten	aantal reacties	gemiddeld gebruik	σ
1ste	1100	8	0.75	2
2de	900	8	2.1	3
3de	700	7	3.6	4
4de	600	7	5.2	5
	3300	30	m	3.6

1. Geef de schatting voor m (gegeven all data).
2. Bepaal de variantie van de schatter indien je de informatie over het studiejaar wil gebruiken.
3. Bepaal de variantie van de schatter indien het studiejaar niet gebruikt wordt. Vergelijk de twee varianties met een toelichting.

Statistiek: Oefeningen hoodstuk 6

1 Methode van de momenten en maximum ikelihood

Oefening 1 (**) Een steekproef bevat n metingen $(X_1, ..., X_n)$ van een variabele X die verdeeld is volgens een waarschijnlijkheidsdichtheid:

$$f(x;\theta,\gamma) = \begin{cases} \frac{1}{\theta}e^{-\frac{x-\gamma}{\theta}} & x \geq \gamma \\ 0 & \text{elders} \end{cases} \quad \text{met} \quad \begin{matrix} \gamma \in \mathbb{R} \\ \theta \in \mathbb{R}_0^+ \end{matrix} \tag{1}$$

Gebruik de methode van de momenten om een schatter op te stellen voor de parameters γ en θ.

Oefening 2 (**) Een variabele $x \in [0, 2\pi]$ is verdeeld volgens een distributie:

$$P(x;\alpha)\mathrm{d}x = \frac{1}{2\pi}(1 + \alpha \sin x)\mathrm{d}x \quad \alpha \in [-1, 1] \tag{2}$$

1. Gebruik de methode van de momenten (kies het moment van de functie $a(x) = x$) om een schatting te maken van de parameter α en de fout daarop.
2. Probeer dit nu met het $\cos x$ en het $\sin x$ moment. Wat stel je vast?
3. Bepaal α aan de hand van de maximum likelihoodmethode.

Oefening 3 (**) Stel dat je twee onafhankelijke metingen hebt met dezelfde nauwkeurigheid: een van $\sin\theta$ en een van $\cos\theta$. Bepaal de ML schatting voor θ.

2 Kleinste kwadraten

Oefening 4 (**) Een radioactieve bron wordt geobserveerd met een Geigerteller. Om het uur wordt gedurende 1 minuut gemeten. Het aantal tellen dat waargenomen wordt is gegeven in de onderstaande tabel.

Tijd t (uur)	0	1	2	3	4	5	6	7	8
Tellen	997	520	265	127	70	35	16	7	3

Bepaal de halfwaardetijd van de bron zowel met behulp van de kleinste kwadratenmethode als de maximum likelihoodmethode. Wat is de variantie op de schatter $\hat{\tau}$ voor grote N?

Oefening 5 (**) Een wagentje beweegt op een spoor aan een constante snelheid. Je tracht deze snelheid te meten. Het wagentje passeert het punt $d = 0$ op tijdstip $t = 0$. Geef de snelheid, de fout op de snelheid en de χ^2 indien...

1. ...we op welbepaalde tijde de positie van het wagentje bepalen, met een nauwkeurigheid van 2 mm:

Tijd t (s)	1.0	2.0	3.0	4.0	5.0	6.0
Afstand d (mm)	11	19	33	40	49	61

2. ...we op welbepaalde afstanden de tijd bepalen, met een nauwkeurigheid van 0.1 s

Afstand d (mm)	10	20	30	40	50	60
Tijd t (s)	1.1	2.2	2.9	4.1	5.0	5.8

Statistiek: Oefeningen hoodstuk 7

Oefening 1 (**) Bij het bestuderen van het verval van protonen (een bijzonder zeldzaam fenomeen) worden 7 evenementen geobserveerd in 1 000 000 kg waterstof gedurende 1 jaar. Geef het 90% confidentie-interval voor het verwachte aantal vervallen λ, en voor de halfwaardetijd die daarmee overeenstemt.

1. Als je aanneemt dat er geen achtergrond is.
2. Als je aanneemt dat er een achtergrond is van 3 evenementen per jaar, veroorzaakt door andere toevallige processen.
3. Als die achtergrond 8 evenementen per jaar bedraagt.

Oefening 2 (**) Een productielijn maakt elektronische sensoren, die per 5 in dozen worden verpakt. Tijdens de kwaliteitscontrole werden er 100 dozen onderzocht. In de volgende tabel zie je een histogram van het aantal defecte eenheden per doos.

Aantal defecte eenheden	0	1	2	3	4	5
Aantal dozen	31	42	18	6	2	1

Bepaal een 90% confidentie-interval voor de fractie defecte sensoren op de productielijn.

Oefening 3 (*) Uit een steekproef van 20 waarnemingen uit een normale verdeling berekent men het 95% confidentie-interval voor het gemiddelde μ. Dit interval blijkt [17.66, 22.34] te zijn.

1. Wat is in dat geval de 90% confidentielimiet voor μ?
2. Veronderstel dat men de steekproefgrootte verhoogt tot $n = 200$, en dat men voor het steekproefgemiddelde en de steekproefvariantie dezelfde waarden uitkomt. Waaraan is dan de 95% confidentielimiet gelijk?

Statistiek: Oefeningen hoodstuk 8

Oefening 1 (*) De vier waarden $\{3.9, 4.5, 5.5, 6.1\}$ worden getrokken gedurende het nemen van een steekproef uit een verdeling waarvan het gemiddelde 4.9 is. Dan wordt 7.3 getrokken. Is het waarschijnlijk dat dit getal uit dezelfde verdeling komt?

Oefening 2 (*) Om een helderziende te onderzoeken wordt het volgende experiment opgezet. Iemand kiest in een aparte kamer een aantal kaarten uit een boek. De helderziende bevindt zich in een andere kamer, en probeert telkens de kleur van de gekozen kaart te raden.

1. Na 50 kaarten had de helderziende 32 keer gelijk. Is dit een betekenisvol resultaat?
2. Hoeveel keer zou de helderziende correct moeten raden opdat zijn prestatie met 99% confidentie significant is?

Oefening 3 (*) Als we 1000 metingen over 25 bins verdelen, er er wordt aan de binwaarden een curve gefit die de som is van een gauss $\mathcal{G}(x; \mu, \sigma)$ en een vlakke achtergrond B, wat is dan het aantal vrijheidsgraden?

Oefening 4 (**) Er worden 10 temperaturen gemeten, elk met een fout van 0.2 K.

$$\{10.2 \quad 10.4 \quad 9.8 \quad 10.5 \quad 9.9 \quad 9.8 \quad 10.3 \quad 10.1 \quad 10.3 \quad 9.9\}$$

1. Er wordt gesuggereerd dat al deze waarden metingen zijn van dezelfde constante, en dat de verschillen puur statistische fluctuaties zijn. Wat is het aantal vrijheidsgraden en de χ^2? Wat is je conclusie?
2. Verandert je conclusie als men beweert dat deze constante waarde 10.1 K is?

Oefening 5 (**) Twee machines zijn ontworpen om identieke staalkabels te produceren. Een steekproef van $n_1 = 11$ waarnemingen levert voor de eerste machine een gemiddelde dikte van $\bar{x_1} = 50.9033\,\mu\mathrm{m}$ op, met een steekproefvariantie van $s_1^2 = 0.1848\,\mu\mathrm{m}^2$. Voor de tweede machine hebben we een steekproef van $n_2 = 8$ waarnemingen met gemiddelde dikte $\bar{x_2} = 50.2525\,\mu\mathrm{m}$ en steekproefvariantie $s_2^2 = 0.2837\,\mu\mathrm{m}^2$.

1. Kunnen we zeggen dat de twee steekproefvarianties dezelfde zijn?
2. Wat kunnen we zeggen over onze nulhypothese?

Oef 1.2 $\alpha_n = \overline{x^n}$ μ_n n-de centrale moment $m = \overline{x}$

Bewijs : $\mu_3 = \alpha_3 - 3m\alpha_2 + 2m^3$

$$\mu_3 = \frac{\sum_i (x_i - \overline{x})^3}{N} = \frac{1}{N}\sum x_i^3 - 3x_i^2\overline{x} + 3x_i\overline{x}^2 - \overline{x}^3$$

$$= \frac{1}{N}\left(N\overline{x^3} - 3N\,\overline{x^2}\,\overline{x} + N3\,\overline{x}^3 - N\overline{x}^3\right)$$

$$= \overline{x^3} - 3\,\overline{x^2}\,\overline{x} + 2\,\overline{x}^3$$

Oef 1.3 €240 over 10 B → €10 5, 5, 10, 10, 50

R → €20 10, 10, 20, 20, 100

1) gem winst ? $\overline{x}_B = \frac{1}{5}\sum x_i = 16$ ⇒ $\overline{\text{winst}} = €6$

$\overline{x}_R = \frac{1}{5}\sum x_i = 32$ ⇒ $\overline{\text{winst}} = €12$

2) 1 v/d €20 verwisseld → 5 mogelijkheden, gemiddeld verschil :

$\frac{\sum x_{i,B}}{5} = 16$ ⇒ gemiddeld is het totale bedrag in de zode €4 hoger

⇒ $\overline{x}_R = 32 - 4/5$ ⇒ $\overline{\text{winst}} = €11,2$

$x_B = 16 + 4/5$ " $= €6,8$

→ Foute interpret.

2) B → €10 5 5 10 10 20 50

$\overline{x}_B = \frac{1}{6}\sum_i x_i = 16,666$

$\overline{x}_R = \frac{1}{4}\sum_i x_i = 35$

⇒ Beide hoger gem hoger ?

Als je ze allemaal wilt kopen moet je nu maar 60 + 4.20 = €140 geven, ipv. €150

Dus er komt geen geld bij, maar de opdeling is beter voor de spelen.

Oef 4.4

t	0,97	1,39	1,69	2,91	3,24	3,95	4,44	4,47	4,56	5,68	5,75
act	31184	13667	7414	654	337	81	31	29	24	3	2

1) corr coeff $\rho = \frac{\text{cov}(act, t)}{\sigma(act)\sigma(t)} = -0,75$ 5,39

$\overline{t} = 3,75$ $\overline{act} = 4452,25$ 1

$$\text{cov} = \frac{\sum (t_i - \overline{t})(act_i - \overline{act})}{N} = \frac{-135027}{N} \approx -11252,25$$

$\sigma(act) = 3022,54$

$\sigma(t) = 1,6568$

$\Rightarrow$ als tijd stijgt : act $\searrow$

2) Haar is eig Int lineair? In exponentieel? $\Rightarrow$ teke log

OP GRM L_1 L_2 L_3 L_4 L_4

	L_1	L_2	L_3	L_4	L_4
	t	act	$t-\bar{t}$	$act-\overline{act}$	$L_3 \cdot L_4$
Varstats $\Rightarrow$	$\bar{t}$	$\overline{act}$	σ	σ	Σ $\Rightarrow$ deel door N

Voor 2: L_1 en t L_2 ~~ln~~ ln act L_3 ln act

$\bar{t} = 3,75$ $\overline{\ln act} = 4,71$

$\sigma_t = 1,65$ $\sigma = 3,34$

$cov = \frac{-66,44}{12} = 5,53$

$\rho \approx 1,00$ $\Rightarrow$ op logschaal is er wel een $\Rightarrow$ perfecte fit

---

Oef 2.1

$Z \rightarrow 1/10\,000$ altijd V en L $\frac{6}{10}$ D $\frac{1}{5}$ N

$\bar{Z} \rightarrow 9999/10000$ $\frac{3}{100}$ V $\frac{1}{10}$ L $\frac{1}{50}$ D $\frac{1}{20}$ N

onafh sympt

1) $P(Z|V,L,D,N) = \frac{P(V,L,D,N|Z)\,P(Z)}{P(V,L,D,N)} =$

Dit vereist vermenigvuldigen $P(V,L) = P(V \cap L) = P(V)P(L)$

$P(VLDN) = P(VLDN|Z)P(Z) + P(VLDN|\bar{Z})P(\bar{Z})$

$= (1 \cdot 1 \cdot \frac{6}{10} \cdot \frac{1}{5}) \cdot \frac{1}{10\,000} + (\frac{3}{100} \cdot \frac{1}{10} \cdot \frac{1}{50} \cdot \frac{1}{20}) \cdot \frac{9999}{10\,000}$

$\Rightarrow P(Z|VLDN) = \frac{6/500\,000}{6/500000 + 3 \cdot 9999/10^{10}} = 0,8$

2) $P(VLD) = P(VLD\bar{N}|Z)P(Z) + P(VLD\bar{N}|\bar{Z})P(\bar{Z})$

$= 1,05 \cdot 10^{-4} = \frac{105}{1\,000\,000}$

Oef 2.2 $R \rightarrow 1\,1\,1\,2\,2\,3$ $B \rightarrow 1\,2\,2\,3\,3\,4$

1) $\langle R \rangle = \frac{10}{6}$ $\langle B \rangle = \frac{15}{6}$

	1	2	3	4
P(R)	1/2	1/3	1/6	0
P(B)	1/6	1/3	1/3	1/6

2) $\langle S \rangle = \sum P(S)\,S = \frac{25}{6}$

	2	3	4	5	6	7
P(S)	1/12	2/9	11/36	9/36	2/18	1/36

3) onafh $\Rightarrow$ $P(R,B) = P(R)P(B)$

B \ R	1	2	3
1	1/12	1/18	1/36
2	1/6	1/9	1/18
3	1/6	1/9	1/18
4	1/12	1/18	1/36

4) $P(M)$?

i	1	2	3	4
$P(M)$	$1/12$	$1/3$	$\frac{15}{36}$	$1/6$

Uit 3 ! Zie zwart

5) $P(X)$

	1	2	3	4	5	6	7	8	9	10	11	12
$P(\Sigma)$	$\frac{1}{12}$	$\frac{19}{108}$	$5/24$								$1/216$	

$\frac{1}{1296}$

Max? $R = 3$, $B = 3 \times 4$

Min? $R1$, $B = 1$

~~$P(X) = P(\Sigma|R)P(R)$~~

$P(X,R) = P(X|R)P(R)$ dit is de conditionele, we willen de marginale!

$$P(X) = \sum_{R=1}^{4} P(X,R)\,P(R)$$

~~$P(1) = \frac{1}{2} \cdot \frac{1}{6}$ $P(2) =$~~

$R = 1 \to X = B \Rightarrow P(X) = P(B) \cdot P(R)$

$R = 2 \to X = B + B \quad = P(B+B) \cdot P(R)$

$R = 3 \to X = B + B + B \quad = P(B+B+B) \cdot P(R)$

	1	2	3	4	5	6	7	8
$B+B$	0	$1/36$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{5}{18}$	$\frac{2}{9}$	$\frac{1}{9}$	$1/36$

	1	2	3	4	5	6	7	8	9	10	11	12
$B+B+B$	0	0	$1/216$								$1/36$	$1/216$

$P(1) = \frac{1}{2} \cdot \frac{1}{6}$ $\quad P_2 = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{36}$ $\quad P(3) = \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{9} + \frac{1}{6} \cdot \frac{1}{216} = 5/24$

$P(11) = \frac{1}{6} \cdot \frac{1}{36}$ $\quad P(12) \cdot \frac{1}{6} \cdot \frac{1}{216}$

Opg 2.3 $\quad \alpha > 0 \quad x > 1$

$P(X > x) = x^{-\alpha}$ Dit is al een soort cumulatieve!

$$= \int_x^{+\infty} f(x)\,dx = x^{-\alpha}$$

$F(\infty) - F(x) = x^{-\alpha}$ als $F(\infty) \neq 0$ is de functie onbegrensd

$$\Rightarrow f(x) = \frac{d}{dx}(-x^{-\alpha}) = \alpha x^{-\alpha-1}$$

$$\langle x \rangle = \int_1^{+\infty} x\,\alpha x^{-\alpha-1}\,dx$$ ~~waarom? begrensd als $\alpha + 1 > 0$~~ $\alpha > -1$

$$= \int_1^{+\infty} \alpha x^{-\alpha}\,dx = \left[\frac{\alpha}{-\alpha+1} x^{-\alpha+1}\right]_1^{+\infty} = \frac{\alpha}{\alpha-1} \quad \text{als } \alpha > 1$$

$$\langle x^2 \rangle = \int_1^{+\infty} \alpha x^{-\alpha+1}\,dx = \left[\frac{\alpha}{-\alpha+2} x^{-\alpha+2}\right]_1^{+\infty} = \frac{+\alpha}{\alpha-2} \quad \text{als } \alpha > 2$$

$$\Rightarrow \text{als } \alpha > 0 \quad V(x) = \frac{\alpha}{\alpha-2} - \left(\frac{\alpha}{\alpha-1}\right)^2$$

Oef 2.4 $P(X=k_1, Y=k_2) = \begin{cases} 1/3 & (k_1,k_2) \in [(0,0),(1,1),(1,-1)] \\ 0 & \notin \end{cases}$

$\rho_{XY} =$

$\langle X \rangle = 2/3$
$\langle Y \rangle = 0$

$cov(X,Y) = \langle XY \rangle - \langle X \rangle \langle Y \rangle$

$\langle XY \rangle = \sum x_i y_i P(x,y) = \frac{1}{3}(0 + 1 - 1)$

$\Rightarrow cov = 0$

Maar $P(X,Y) \neq P(X)P(Y)$ wat $P(X) = \begin{cases} 1/3 & x=0 \\ 2/3 & x=1 \end{cases}$

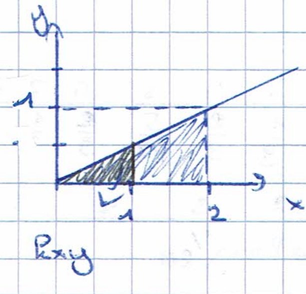
$P(Y) = \begin{cases} \frac{1}{3} & y = \{-1, 0, 1\} \end{cases}$

$\Rightarrow$ ~~$P(x)P(y) =$~~ $\{$

Oef 2.5

$X \in [0,2]$ $Y \in [0,1]$

$f_{XY}(x,y) = \begin{cases} kxy & 2y < x \\ 0 & \text{elders} \end{cases}$



1) k? door de normering!

$$\int_0^2 dx \int_0^{x/2} kxy\, dy = 1$$

$$\int_0^2 kx \left[y^2/2\right]_0^{x/2} dx = \int_0^2 kx^3/8\, dx = \left[\frac{kx^4}{32}\right]_0^2 = \frac{k}{2} \Rightarrow k = 2$$

2) $P(X<1, Y<0{,}5)$

$$= \int_0^1 dx \int_0^{x/2} 2xy\, dy = \left[\frac{2x^4}{32}\right]_0^1 = 1/16$$

3) Marg $f_x = \int_0^{x/2} 2xy\, dy = \frac{x^3}{4}$ $\qquad f_y = \int_{2y}^2 2xy\, dx = 2y\left[x^2/2\right]_{2y}^2$

$= 4y - 4y^3$

Cond $f_x(x|y) = \frac{f_{xy}}{f_y} = \frac{2xy}{4y-4y^3} = \frac{1}{2}x\left(\frac{1}{1-y^2}\right)$

$\Rightarrow$ hoe dichter y bij 1, hoe groter de kan (en $y=1$ kan nt, daar is maar 1pt)

$$f_y(y|x) = \frac{f_{xy}}{f_x} = \frac{2xy}{x^3/4} = 8y/x^2$$

$E(Y|x) = \int_0^{x/2} y\, 2xy\, dy = 2x\left[y^3/3\right]_0^{x/2} = \frac{x^4}{12}$ → hoe grotere x, hoe grotere je Y verwacht.

vast

a hoge macht omdat a véél meer pte zijn bij grotere x.

Oef 3.1

Bewijs de geheugenloosheid

$$P(X > x_1+x_2 \mid X > x_2) = \frac{P(X > x_2 \mid X > x_1+x_2)\, P(X > x_1+x_2)}{P(X > x_2)}$$

$$= \frac{P(X > x_1+x_2)}{P(X > x_2)}$$ → dit zijn cumulatieve, maar van $+\infty$

$$= \frac{1 - F(x_1+x_2, \tau)}{1 - F(x_2, \tau)} = \frac{e^{-(x_1+x_2)/\tau}}{e^{-x_2/\tau}} = e^{-x_1/\tau}$$

$$= 1 - F(x_1, \tau)$$

$$= P(X > x_1)$$

Oef 3.2

$B(p=0,5 \ ; \ n=16)$ Bereken $P(X \leq 5)$

1) Via Binom : neem gewoon de som

$$\sum_{i=0}^{5} \frac{16!}{i!(16-i)!}(0,5)^i(0,5)^{16-i} = \sum_{i=0}^{5} \frac{16!}{i!(16-i)!} \cdot 0,5^{16}$$

$$= 0,5^{16} + 16 \cdot 0,5^{16} + 120 \cdot 0,5^{16} + 560 \cdot 0,5^{16} + 1820 \cdot 0,5^{16} + 4368 \cdot 0,5^{16}$$

$= 10,5\ \%$

---

2) via poisson : $\lambda = np$ ~~$+2\ 3p2+p$~~

$\lambda = 8$

$$\Rightarrow P(X \leq 5) = \sum_{i=0}^{5} \frac{8^i e^{-8}}{i!} = 19,1\ \%$$

---

3) via Gauss : $\mu = np$ $\sigma = \sqrt{np(1-p)}$ $\mu = 8$ $\sigma = 2$

$$P(X \leq 5) = -\int_{-\infty}^{0} G\,dx + \int_{-\infty}^{5} G\,dx$$

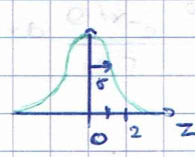
grenzen? $z = \frac{x-\mu}{\sigma}$

$\Rightarrow x = 0 \rightsquigarrow z = -4$

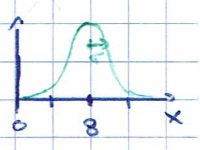
$x = 5 \rightsquigarrow z = -3/2$

Dus zoek nu voor deze z-waarden in de tabel de cumulatieve

$$\int_{-\infty}^{-4} N\,dz \approx 0 \qquad \int_{-\infty}^{-3/2} N\,dz = 1 - \int_{3/2}^{+\infty} = 6,68\ \%$$



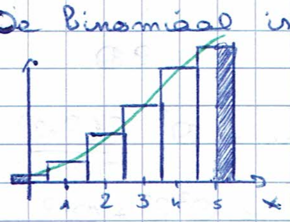
←



---

4) Continuïteitscorrectie?

De Binomiaal is discreet! Intervallen zijn gecentreerd rond de gehele getallen.



$\Rightarrow$ Als je Gauss neemt, doe vanaf $-0,5$ tot $5,5$

$$P(X \geq 11) = \int_{-\infty}^{5,5} G\,dx - \int_{-\infty}^{-0,5} G\,dx = \int_{-\infty}^{-1,25} N\,dz - \int_{-\infty}^{-4,25} N\,dz = 89,44\ \%$$

$\Rightarrow P(X \leq 5) = 10,56\ \%$

Oef 3 $\quad FWHM = 2\sigma\sqrt{2\ln 2} = 2{,}35\,\sigma$ voor Gauss

Op halve hoogte van max? Wat is max hoogte?

$$x=\mu \quad G(\mu) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{(0)^2/2\sigma^2} = \frac{1}{\sqrt{2\pi}\,\sigma}$$

$\Rightarrow$ halve hoogte $\dfrac{1}{2\sigma\sqrt{\pi 2}}$ wanneer is $G(x)$ hieraan gelijk?

$$\frac{1}{2\sigma\sqrt{\pi 2}} = \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left[-(x-\mu)^2/2\sigma^2\right]$$

$$\Rightarrow 1 = 2\exp[\ldots]$$

$$\Rightarrow 0 = \ln(2) + \left(-(x-\mu)^2/2\sigma^2\right)$$

$$\Rightarrow 2\sigma^2\ln(2) = (x-\mu)^2$$

$\Rightarrow x-\mu = \sqrt{2\ln(2)}\,\sigma$ is de helft v/d breedte $\Rightarrow$ gevr.

Oef 3.4

$\lambda = 1/_{\text{minuut}} \quad P(\text{ligt}) = 1\%$

1) kans dat hij er na 60 auto's nog staat

$$\Rightarrow \text{Bin}(0\,;0{,}01;60) = 1\cdot(0{,}01)^0(0{,}99)^{60} = 54{,}7\%$$

2) na 1 uur? $\lambda = 60 \rightarrow$ maar evenement dat hij meegenomen is $\rightarrow \frac{60}{100}$

$$\Rightarrow P(0;0{,}6) = \frac{0{,}6^0\,e^{-0{,}6}}{0!} = 54{,}9\%$$

Oef 3.5 Gammafunctie

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

1) $$\Gamma(x+1) = \int_0^{+\infty} e^{-t}\,t^x\,dt = \underbrace{\left[-e^{-t}\,t^x\right]_0^{+\infty}}_{0} + \int_0^{+\infty} e^{-t}\,x\,t^{x-1}\,dt = x\,\Gamma(x-1)$$

---

2) $$\langle X\rangle = \int_0^{+\infty}\frac{x^k\,e^{-x/\theta}}{\theta^k\,\Gamma(k)}\,dx = \frac{1}{\theta^k\,\Gamma(k)}\int_0^{+\infty} x^k\,e^{-x/\theta}\,dx \qquad \begin{matrix} u = x/\theta \\ x = \theta u\end{matrix} \qquad du = \frac{dx}{\theta}$$

$$= \frac{\theta^{k+1}}{\theta^k\,\Gamma(k)}\int_0^{+\infty} u^k\,e^{-u}\,du \quad \rightarrow \text{herken } \Gamma?$$

$$= \frac{\theta}{\Gamma(k)}\int_0^{+\infty} e^{-u}\,u^{(k+1)-1}\,du = \theta\,\frac{\Gamma(k+1)}{\Gamma(k)}$$ als $k\in\mathbb{N}$, dan $= k\theta$

Variantie? $$\langle x^2\rangle = \int_0^{+\infty}\frac{x^{k+1}\,e^{-x/\theta}}{\theta^k\,\Gamma(k)}\,dx = \frac{\theta^2}{\Gamma(k)}\int_0^{+\infty} u^{k+1}\,e^{-u}\,du = \theta^2\,\frac{\Gamma(k+2)}{\Gamma(k)} = \theta^2k^2 + \theta^2 k$$

$$\Rightarrow V(x) = \langle x^2\rangle - \langle x\rangle^2 = \theta^2 k$$

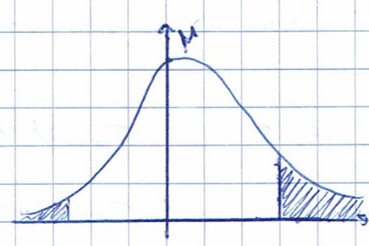
3) exponentiële ($P(x) = \frac{e^{-x/\xi}}{\xi}$) is speciaal geval?

als $k = 1$ en $\theta = \xi$ $\Rightarrow f(x, 1, \xi) = \frac{x^0 e^{-x/\xi}}{\theta^1 \underbrace{\Gamma(1)}_{=0!}}$

$\Rightarrow \langle x \rangle = \xi \qquad V(x) = \xi^2$

Oef 3.6

Normaal verdeeld $\qquad$ 10% : $> 10{,}256\,\Omega$
$\qquad\qquad$ 5% : $< 9{,}671$



gevr: μ en σ

opl: zoek in tabellen de z-waarde voor 10% & 5% (eenzijdig)

Bovenste 10% : $1{,}28 = z$ $\qquad$ Laagste 5% : $1{,}645 = z$

$z = \frac{x-\mu}{\sigma}$ $\qquad \Rightarrow \begin{cases} 1{,}28\,\sigma = 10{,}256 - \mu \\ 1{,}645\,\sigma = \mu - 9{,}671 \end{cases}$ $\rightarrow$ want langs links

$\Rightarrow -1{,}28\,\sigma + 10{,}256 = + 1{,}645\,\sigma + 9{,}671$

$2{,}925\,\sigma = 0{,}585 \Rightarrow \sigma = 0{,}2\,\Omega$

$\Rightarrow \mu = 10\,\Omega$

Oef 4.1

$\sin i = n \sin 2$

$i\ (\pm 1)^\circ$	$2\ (\pm 1)^\circ$
20	13
40	23,5

1) geef de brekingsindex

Foutenprop? $\qquad n = \frac{\sin i}{\sin 2} \Rightarrow \left(\frac{\sigma_n}{n}\right)^2 = \left(\frac{\sigma(\sin i)}{\sin i}\right)^2 + \left(\frac{\sigma(\sin 2)}{\sin 2}\right)^2$

$\sigma^2(\sin i) = \left(\frac{\partial \sin i}{\partial i}\right)^2 \sigma_i^2 \Rightarrow \left(\frac{\sigma_n}{n}\right)^2 = \left(\frac{\cos i}{\sin i}\,\sigma_i\right)^2 + \left(\frac{\cos^2 2}{\sin 2}\,\sigma_2\right)^2$

$= \cos i^2\, \sigma_i^2$

$\sigma_n^2 = \left(\frac{\cos^2 i}{\sin(2)}\right.$

Kan op 2 manieren ♀

1) Via fractionele foute (want het is een quotiënt)

2) Via gewone foutepropagatie

Fractionele fouten zijn niets 'nieuws', gewoon een simpelere vorm v/d algemene ♀ (In sommige gevallen)

$V(n) = \left(\frac{\partial n}{\partial i}\right)^2 \sigma_i^2 + \left(\frac{\partial n}{\partial 2}\right)^2 \sigma_2^2$ $\qquad$ (2 en i zijn onafh.)

$$\Rightarrow V(n) = \frac{\cos^2(i)}{\sin^2(2)}\sigma_i^2 + \left(\frac{\sin(i)\cos(2)}{\sin^2(2)}\right)^2\sigma_2^2$$

(via fract: $\frac{V_n}{n^2} = \frac{V(\sin i)}{(\sin i)^2} + \frac{V(\sin 2)}{(\sin 2)^2}$ $\Rightarrow V_n = n^2\frac{\cos^2(i)}{\sin^2(i)}\sigma^2(i) + \frac{n^2\cos^2(2)}{\sin^2(2)}\sigma^2_{(2)}$

$V(\sin i) = \cos^2(i)V(i)$
$V(\sin 2) = \cos^2(2)V(2)$

$$V_n = \frac{\cos^2(i)}{\sin^2(2)}\sigma_i^2 + \frac{\sin^2(i)\cos^2(2)}{\sin^4(2)}\sigma^2_{(2)}$$

$\boxed{\text{Hoeken altijd in radialen}}$

$n_{(1)} = 1,52$ $\quad \sigma_1 = 0,136$

$n_{(2)} = 1,61$ $\quad \sigma_2 = 0,064$

$\Rightarrow$ volgens de eerste waarde wel compatibel, volgens 2° niet

Extra: H8 toepassen: 2 waarden compatibel?

$n_2 - n_1$ compatibel met 0

$\Rightarrow \mu = n_2 - n_1 = 0,09$

$\sigma = \sqrt{\sigma_i^2 + \sigma_2^2} = 0,150$ $\quad \Rightarrow$ 0 ligt zelfs binnen 1σ
$\Rightarrow H_0\,(n_1 = n_2)$ niet verwerpen!

Oef 4.2

De 2 waarden hebben dezelfde (ongekende) fout $\Rightarrow$ gewoon gemiddelde

~~tange bavensdaags~~ $\Rightarrow$ Poisson verdeeld! Dus elk vid

$\bar{\lambda} = \frac{\lambda_1 + \lambda_2}{2} = 406$ $\quad \cancel{V(\bar{\lambda}) = \frac{V(\lambda)}{N}}$ $\quad V(\lambda) = V\left(\frac{\lambda_1+\lambda_2}{2}\right) = \frac{1}{4}V(\lambda_1) + \frac{1}{4}V(\lambda_2)$

$\Rightarrow$ Poisson verdeeld

$V(\lambda) = \frac{1}{4}\lambda_1 + \frac{1}{4}\lambda_2 = 203$

$\sigma(\lambda) = 14,24$

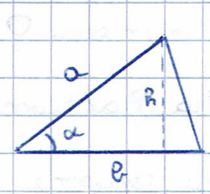
NE interessant met zicht op volgende hoofdstukken

Oef 4.3

Omtrek en oppervlakte v. een driehoek waarvan α, a, b gekend zijn (ongecorrelleerd)

Omtrek? Begin met de cosinusregel!

$c^2 = a^2 + b^2 - 2ab\cos(\alpha)$



$h = a\sin(\alpha)$

$L = a + b + \sqrt{a^2 + b^2 - 2ab\cos(\alpha)}$ $\quad A = \frac{b\,a\sin(\alpha)}{2}$

1) $V(L) = \left(\frac{\partial L}{\partial a}\right)^2 V(a) + \left(\frac{\partial L}{\partial b}\right)^2 V(b) + \left(\frac{\partial L}{\partial \alpha}\right)^2 V(\alpha)$ (ongecorreleerd)

$= \left[1 + \frac{1}{2}(a^2+b^2-2ab\cos\alpha)^{-1/2}(2a - 2b\cos\alpha)\right]^2 V(a) + \left[1 + \frac{1}{2}(a^2+b^2-2ab\cos(\alpha))^{-1/2}(2b - 2a\cos\alpha)\right]^2 V(b)$

$+ \left[\frac{1}{2}(a^2+b^2-2ab\cos(\alpha))^{-1/2}(2ab\sin(\alpha))\right]^2 V(\alpha)$

$$V(A) = \left(\frac{\partial A}{\partial a}\right)^2 V(a) + \left(\frac{\partial A}{\partial b}\right)^2 V(b) + \left(\frac{\partial A}{\partial \alpha}\right) V(\alpha)$$

$$= \left[\frac{b\sin(\alpha)}{2}\right]^2 V(a) + \left[\frac{a\sin(\alpha)}{2}\right]^2 V(b) + \left[\frac{ab\cos(\alpha)}{2}\right]^2 V(\alpha)$$

$$\mathrm{cov}(L,A) = \frac{\partial A}{\partial a}\frac{\partial L}{\partial a} V(a) + \frac{\partial A}{\partial b}\frac{\partial L}{\partial b} V(a) + \frac{\partial A}{\partial \alpha}\frac{\partial L}{\partial \alpha} V(\alpha)$$ (door de afwezigheid van correlaties)

2) Toegepast op de waarden geeft dit

$L = 303,817 \pm 0,093$

$A = 3862,4 \pm 2,6$

$\rho = 0,992 \rightarrow$ Reel sterk gecorrelleerd : hoe groter ~~of ot hoe groter~~ beide de ene, hoe groter de andere logischerwijs ook is.

3) $V(a) \rightarrow V(a) + s^2$
$V(b) \rightarrow V(b) + s^2$
$\mathrm{cov}(a,b) = s^2$

$\Rightarrow$ De fouten op L en A nemen toe $\quad \sigma_L = 1,4$
$\sigma_A = 42$

en er zijn extra termen in de $\mathrm{cov}(L,A) \Rightarrow$ vergroot ook : $\rho = 0,99997$

Opg 5.1

Binomiaal : N_S succes N_F falen

$\hat{p} = \frac{N_S}{N_S + N_F}$ consis en onbev. ? $\lim_{+\infty} \hat{p} = \lim_{+\infty} \frac{N_S}{N_{Tot}} = p$ volgens de wet v/d grote getallen

en $\langle \hat{p} \rangle = \frac{\langle N_S \rangle}{N_{Tot}} = \sum_{N_S=0}^{N_{Tot}} \frac{N_S}{N_{tot}} \cdot p = \frac{1}{N_{tot}} \sum_0^{N_{tot}} \frac{N_S\, N_{tot}!}{N_S!\,N_F!} p^{N_S} p^{N_{tot}-N_S}$

$= \frac{1}{N_{tot}} N_{tot}\, p \sum \frac{(N_{tot}-1)!}{(N_S-1)(N_{tot}-N_S)} p^{N_S-1} p^{N_{tot}-N_S}$

$= p$

Opg 5.2

Poisson verdeling

1) $(\bar{x})^2$ geen onbevooroordeelde schatter?

$\langle (\bar{x})^2 \rangle = \langle (\frac{1}{N}\sum_{i=1}^{N} x_i)^2 \rangle = \frac{1}{N^2} \langle (\sum_i^N x_i)(\sum_j^N x_j) \rangle = \frac{1}{N^2} \langle \sum_{i=j} x_i^2 \rangle + \frac{1}{N^2} \langle \sum_{i \neq j} x_i x_j \rangle$

$= \frac{1}{N^2} \sum_i \langle x_i^2 \rangle + \frac{1}{N^2} \sum_{i,j} \langle x_i \rangle \langle x_j \rangle$

Want de verwachtingswaarde van x_i zijn $= \langle X \rangle = \lambda$ $\rightarrow$ $= \frac{1}{N} \lambda^2 + \frac{N(N-1)}{N^2} \lambda^2$ Niet

Wel $= \frac{1}{N^2} \sum_i \langle x_i^2 \rangle + \frac{1}{N^2} \sum_{i \neq j} \langle x_i \rangle \langle x_j \rangle$ $=$

$= \frac{1}{N^2} \sum \langle x^2 \rangle + \frac{1}{N^2} \sum \langle x_i \rangle \langle x_j \rangle$ en $V(x) = \lambda = \langle x^2 \rangle - \langle x \rangle^2$
$\Rightarrow \langle x^2 \rangle = \lambda + \langle x \rangle^2$

$$= \frac{N(\lambda + \lambda^2)}{N^2} + \frac{(N-1)N(\lambda^2)}{N^2} = \frac{\lambda}{N} + \frac{\lambda^2}{N} + \lambda^2 - \frac{\lambda^2}{N} = \frac{\lambda}{N} + \lambda^2 \neq \lambda^2$$

=> Bevooroordeeld

$$\boxed{\langle x^2 \rangle \neq \langle x \rangle^2 \quad \text{NOOIT ZOMAAR}}$$

2) De bias wegwerken? $\lambda = \bar{X}$ => $\widehat{\lambda^2} = \bar{X}^2 - \frac{\bar{X}}{N}$

Oef 5.3

20.0 19,7 20,6 18,5 21,2 20,8 20,7

1) $\hat{\mu} = \bar{X} = 20,21$ $V(\bar{x}) = \frac{V(x)}{N}$ => $\sigma(\bar{x}) = \frac{0,8}{\sqrt{7}} = 0,30$

2) $\hat{V}(x)$ met gekend gemiddelde: $\frac{1}{N}\sum_i^N (x_i - \mu)^2 = 0,75$

Int over volledige? => $\hat{\sigma} = 0,87$

Fout hierop? $\hat{V}[\widehat{\sigma(x)}] = \frac{1}{4N\sigma^2}\left[\langle (x - \langle x \rangle)^4 \rangle - \langle (x - \langle x \rangle)^2 \rangle^2\right]$

$$= \frac{1}{4N\sigma^2}\left[\overline{(x-\mu)^4} - \overline{(x-\mu)^2}^2\right]$$

$$= \frac{1}{28 \cdot 0,75}\left[1,1319 - 0,5667\right] = 0,0268 \rightsquigarrow \sigma(\hat{\sigma}(x)) = 0,164$$

3) $\widehat{V(x)}$ met ongekend gemiddelde => correctie van Bessel $\frac{1}{N-1}$ ipv $\frac{1}{N}$

Voor de rest volledig analoog, $\mu \leftrightarrow \bar{x}$

$s = 0,841$, $\sigma(s(x)) = 0,214$

4) $\hat{\mu} = \bar{x} = 20,21$ nog steeds, maar nu $V(\bar{x}) = \frac{s^2}{N} = 0,318$

Oef 5.4

1) $\hat{m} = \frac{1}{300}\sum_i^4 N_i \cdot m_i = 2,813$

2) Bepaal variatie, indien je het jaar gebruikt => Stratified sampling

5.16: ~~$V = \sum_{i=1}^4 f_i^2 \frac{V_i}{m_i} = \frac{1}{30^2}\sum_{i=1}^4 m_i V_i = \frac{1}{300}$~~ Gebruik de formule in de opgave

$V(\hat{m}) = \underbrace{\sum_{i=1}^4 f_i^2 V_i}_{\text{Ts opgave}} = \sum_{i=1}^4 \left(\frac{N_i}{N_{tot}}\right)^2 \frac{N_i - n_i}{N_i - 1} \frac{\sigma_i^2}{n_i}$

N_i: totaal aantal stud. v/d deelgroep
n_i: aantal in je steekproef " "
σ_i: je steekproef afwijking " "

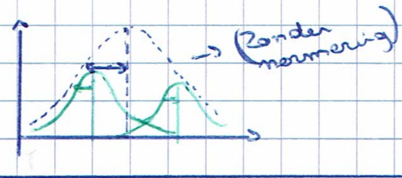
$$= \left(\frac{1100}{3300}\right)^2 \frac{1092}{1099} \frac{2^2}{8} + \ldots$$

$$= 0,055 + 0,083 + 0,102 + 0,119 = 0,357$$

31 Wat als je de stratified sampling niet toepast?

$$V(\hat{m}) = \frac{N-n}{N-1}\,\frac{\sigma^2}{n} = \frac{3300-30}{3299}\,\frac{3{,}6}{30} = 0{,}428$$

Deze is duidelijk groter dan de $V(\hat{m})$ uit 21. Dit verschil is afkomstig uit het feit dat je maar 1 verdeling veronderstelt



Oef 6.1

$$f(x,\Theta,\gamma) = \begin{cases} 1/\Theta \cdot \exp\left[-\frac{(x-\gamma)}{\Theta}\right] & x \geq \gamma \\ 0 & \text{elders} \end{cases} \qquad \gamma \in \mathbb{R},\ \Theta \in \mathbb{R}^+_0$$

Schatten via de methode v/d momenten

Stappenplan:
1. Kies functies g_j, als je de verdeling herkent zijn de momenten zeker een goede keuze (dus bv $g_1 = x$ $g_2 = x^2$)
2. Bepaal, via $\langle\cdot\rangle = \int \cdot\, P\,dx$ welke vorm je g_j kan aannemen
3. Vul je theoretische momenten in waar mogelijk
4. $n \times n$ stelsel op te lossen
5. $\langle\cdot\rangle = \overline{g_i}$

Dit is een exponentiële verdeling $\Rightarrow \langle x\rangle = \Theta$, $\langle x^2\rangle = 2\Theta^2$ → vind je bv uit $V = \Theta^2$

$$M_1 = \int_\gamma^{+\infty} \frac{x}{\Theta}\exp[-(x-\gamma)/\Theta]\,dx \qquad y = x-\gamma \quad dy = dx \quad x=\gamma \Rightarrow y = 0$$

$$= \int_0^{+\infty} \frac{y+\gamma}{\Theta}\exp[-y/\Theta]\,dy = \Theta + \underbrace{1}_{\text{normering}}\cdot\gamma$$

$$M_2 = \int_\gamma^{+\infty} \frac{x^2}{\Theta}\exp[-(x-\gamma)/\Theta]\,dx = \int_0^{+\infty} \frac{y^2+2y\gamma+\gamma^2}{\Theta}\exp[-y/\Theta]\,dy$$

$$= 2\Theta^2 + 2\gamma\Theta + \gamma^2\cdot 1$$

$$\Rightarrow \begin{cases} M_1 = \Theta+\gamma \\ M_2 = 2\Theta^2 + 2\gamma\Theta + \gamma^2 \end{cases} \Rightarrow \begin{cases} \gamma = M_1 - \Theta \\ M_2 = 2\Theta^2 + 2M_1\Theta - 2\Theta^2 + \Theta^2 - 2M_1\Theta + M_1^2 \end{cases}$$

$$\Rightarrow \begin{cases} \gamma = M_1 - \Theta \\ \Theta = \sqrt{M_2 - M_1^2} \end{cases}$$

Nu pas doe je de cruciale stap, je schatting: $M_1 = \overline{x}$, $M_2 = \overline{x^2}$

$$\Rightarrow \quad \hat{\Theta} = \sqrt{\overline{x^2} - \overline{x}^2} \qquad \hat{\gamma} = \overline{x} - \sqrt{\overline{x^2} - \overline{x}^2}$$

Oef 6.2 $\quad x \in [0, 2\pi] \quad P(x,\alpha)\,dx = \frac{1}{2\pi}(1+\alpha\sin x)\,dx \quad \alpha \in [-1,1]$

1) We gebruiken $g_1 = x$

$$\to M_1 = \int_0^{2\pi} \frac{x}{2\pi}(1+\alpha\sin x)\,dx = \left[\frac{x^2}{4\pi}\right]_0^{2\pi} + \int_0^{2\pi} \frac{\alpha x}{2\pi}\sin x\,dx$$

$$= \pi + \left[-\frac{\alpha x}{2\pi}\cos(x)\right]_0^{2\pi} + \int_0^{2\pi}\cos(x)\,dx$$

$$= \pi - \alpha$$

$$M_1 = \pi - \alpha \Rightarrow \hat{\alpha} = \pi - \bar{x} \qquad V(\hat{\alpha}) = \left(\frac{d\alpha}{dx}\right)^2 V(\bar{x}) = 1 \cdot \frac{V(x)}{N}$$

2) $g_1 = \cos x$

$$\Rightarrow M_{\cos} = \int_0^{2\pi} \frac{1}{2\pi}\cos x\,(1+\alpha\sin x)\,dx = 0 + \int_0^{2\pi}\frac{\alpha}{2\pi}\cos x \sin x\,dx \qquad \begin{matrix} u = \sin x \\ du = \cos x\,dx \end{matrix}$$

$$= \frac{\alpha}{2\pi}\int_0^0 \ldots \Rightarrow \text{nt bruikbaar, geeft } 0\,!$$

$g_1 = \sin x$

$$\Rightarrow M_{\sin} = \int_0^{2\pi}\frac{\sin x}{2\pi}(1+\alpha\sin(x))\,dx = 0 + \int_0^{2\pi}\frac{\alpha}{2\pi}\sin^2(x)\,dx = \int_0^{2\pi}\frac{\alpha}{2\pi}\left(\frac{1-\cos(2x)}{2}\right)dx$$

$$= \frac{\alpha}{2} - \frac{\alpha}{4\pi}\int_0^{4\pi}\cos(u)\,du = \frac{\alpha}{2}$$

$$\Rightarrow M_{\sin} = \frac{\alpha}{2}$$

$$\Rightarrow \hat{\alpha} = 2\,\overline{\sin(x)} \qquad V(\hat{\alpha}) = \left(\frac{d\hat{\alpha}}{d\sin(x)}\right)^2 V(\overline{\sin(x)})$$

$$= \frac{4\,V(\sin(x))}{N} = \frac{4\cos^2(x)\,V(x)}{N}$$

3) via LLHM

$$\mathcal{L}(\alpha, x_i) = \prod_{i=1} P(\alpha, x_i) \Rightarrow \ln\mathcal{L} = -N\ln(2\pi) + \sum_i^N \ln(1+\alpha\sin x_i)$$

de index i staat niet voor de i-de parameter, maar voor de i-de meting v/d steekproef !

$$\Rightarrow \frac{\partial \ln\mathcal{L}}{\partial\alpha} = \sum_i (1+\alpha\sin(x_i))^{-1}\sin(x_i) \qquad \Rightarrow \text{De } \alpha \text{ waarvoor dit} = 0, \text{ is } \hat{\alpha}.$$

(Je kan dit normaal verder uitwerken, maar hier heb ik geen zin in.x)
(Wordt in de les ook niet gedaan)
(Doe numeriek of zo)

Opg 6.4

t (u)	0	1	2	3	4	5	6	7	8
a	997	520	265	127	70	35	16	7	3

De activiteit wordt door een exponentiële verdeling beschreven

$$a(t) = A \exp[-t/\tau]$$

Iets wat we hier handig kunnen toepassen is de theorie rond het fitten van een rechte:

$$\ln(a(t)) = \ln(A) - t/\tau \qquad \Rightarrow \quad m = -1/\tau$$

Dus door m te fitten (lees: schatten) hebben we ook een schatter voor τ

$$(6.20) \rightarrow \hat{m} = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} \qquad \text{Dus} \quad \hat{m} = \frac{\overline{\ln(y)t} - \overline{\ln(y)}\,\bar{t}}{\overline{t^2} - \bar{t}^2}$$

Maar let op! Fout op de a_i is dezelfde, maar op $\ln(a_i)$ niet!

$$V(\ln(a_i)) = \frac{1}{a_i^2} V(a_i) \qquad a_i \text{ is tellen: Poisson verdeeld rond z'n waarde} \Rightarrow V(a_i) = a_i$$

$$= 1/a_i$$

We gebruiken dus het gewogen gemiddelde, van ~~de a_i's~~.

$$\overline{t^2} = \frac{1}{8} \Sigma t_i^2 = 25{,}5 \qquad \bar{t} = 4 \qquad V(t) = 7{,}5$$

$$\bar{a} = \frac{\Sigma a_i/(1/a_i)}{\Sigma 1/(1/a_i)} = 665{,}29 \qquad \Sigma a_i = 2040$$

$$\overline{at} = \frac{1}{2040} \Sigma a_i^2 t_i = 238{,}68 \quad \Leftrightarrow$$

$$\Rightarrow \hat{m} =$$

Het gewogen gemiddelde moet op alles toegepast worden! Zie p141: als je $1/\sigma_i^2$ niet mag buiten brengen

$$W = \Sigma 1/\sigma_i^2 = \Sigma a_i = 2040$$

$$\bar{t} = \frac{1}{W} \Sigma t_i/(1/a_i) = 1{,}007$$

$$\overline{t^2} = \frac{1}{W} \Sigma t_i^2 a_i = 2{,}857$$

$$\overline{\ln(y)t} = \frac{1}{W} \Sigma t_i a_i \ln(a_i) = 5{,}027$$

$$\overline{\ln(y)} = \frac{1}{W} \Sigma a_i \ln(a_i) = 6{,}231$$

$$\Rightarrow \hat{m} = -0{,}6789$$
$$\Rightarrow \tau = 1{,}473$$

Variantie? Gebruik (6.24) en (6.21)

$$V(\hat{m}) = \frac{1}{N(\overline{t^2} - \bar{t}^2)} \frac{N}{W} = 0{,}000266 \Rightarrow V(\hat{\tau}) = \left(\frac{d\hat{\tau}}{d\hat{m}}\right)^2 V(\hat{m})$$

$$= \frac{1}{\hat{m}^4} V(\hat{m}) = 0{,}00125$$

$$\Rightarrow \sigma(\hat{\tau}) = 0{,}035$$

Zie volgende pag voor LLHM

Oef 6.3 (Sorry, gemist)

2 onafh. metingen, zelfde σ $\quad \underset{A \pm \sigma}{\underset{\parallel}{\sin\theta}}$ en $\underset{B \pm \sigma}{\underset{\parallel}{\cos\theta}}$

Bepaal ML voor θ

We veronderstellen dat $\sin\theta$ en $\cos\theta$ Gaussisch verdeeld zijn rond A en B (dit kan meestal wel)(Bij benadering)

$$\Rightarrow \mathcal{L}(A,B,\sigma,\theta) = \underbrace{P(\sin\theta = A \pm \sigma ; \theta)}_{\text{Gauss, verdeeld rond } A} \cdot \underbrace{P(\cos\theta = B \pm \sigma ; \theta)}_{\text{Gauss, verdeeld rond } B}$$

$$= \frac{1}{2\pi\sigma^2} \exp[-(\sin\theta - A)^2/2\sigma^2] \exp[-(\cos\theta - B)^2/2\sigma^2]$$

$$\Rightarrow \ln\mathcal{L} = -\ln(2\pi\sigma^2) - \frac{(\sin\theta - A)^2}{2\sigma^2} - \frac{(\cos\theta - B)^2}{2\sigma^2}$$

$$\frac{\partial \ln\mathcal{L}}{\partial\theta} = \frac{1}{2\sigma^2}\left[-2(\sin\theta - A)\cos\theta + 2(\cos\theta - B)\sin\theta\right]$$

$$\Rightarrow \text{stel} = 0 : -\sin\hat\theta\cos\hat\theta + A\cos\hat\theta + \sin\hat\theta\cos\hat\theta - B\sin\hat\theta = 0$$

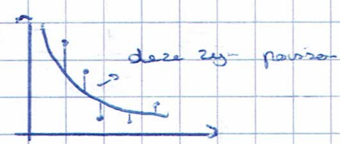
$$\Rightarrow \tan\hat\theta = A/B \quad \Rightarrow \hat\theta = \mathrm{Bgtan}(A/B)$$

Oef 6.4 vervolg

Via de LLHM : $P(a_i \mid \lambda(t, A, \tau))$ want de kans op een bepaalde a_i is verdeeld volgens ee poissonverdeling.

Nu is $\underset{\hookrightarrow \text{ifv. } A, t, \tau}{\lambda = A\exp[-t/\tau]}$ de verwachtingswaarde voor deze poisson

$$\Rightarrow P(a_i, \lambda) = \frac{e^{-\lambda}\lambda^{a_i}}{a_i!}$$



$$\Rightarrow \mathcal{L} = \prod_i P(a_i, \lambda) = \prod_i \frac{e^{-\lambda}\lambda^{a_i}}{a_i!}$$

$$\ln(\mathcal{L}) = \sum_i a_i \ln(\lambda) - \lambda - \ln(a_i!) = \sum_i a_i\left(\frac{-t}{\tau}\right) + a_i \ln(A) - Ae^{-t/\tau} - \ln(a_i!)$$

$$\frac{\partial(\ln(\mathcal{L}))}{\partial\tau} = \sum_i \frac{a_i t}{\tau^2} - Ae^{-t/\tau} \cdot \frac{t_i}{\tau^2} \rightarrow = 0$$

Om $V(\hat\tau)$ te bepalen: voor grote N is MLLH efficiënt $\Rightarrow$ MVB

$$V(\hat\tau) = \frac{1}{\left\langle\left(\frac{d(\ln(\mathcal{L}))}{d\tau}\right)^2\right\rangle}$$

Numeriek : $\hat\tau = 1,471$

Oef 6.5 $\quad v = c^{te} \quad$ gevr: v, $V(v)$ en χ^2

① $\sigma = 2mm$

t(s)	1,0	2,0	3,0	4,0	5,0	6,0
d(mm)	11	19	33	40	49	61

$d = vt$ => Je fit een rechte door de oorsprong

$$\Rightarrow \bar{v} = \frac{\overline{dt}}{\overline{t^2}} \approx 10,099$$

(merk op: niet gewoon $\frac{\bar{d}}{\bar{t}}$! Dit geeft 10,14)

$$V(\hat{v}) = \frac{\sigma_d^2}{N\,\overline{t^2}} \approx 0,203 \text{ mm}$$

$$\chi^2 = \sum \frac{1}{2^2}(d_i - vt_i)^2 = 3,03$$

dit is dus minimaal!
$\approx N/2$, dus redelijk klein!

② $\sigma(t) = 0,1s$

nu is de situatie gedraaid! De theorie gaat uit van een y met fout,

Bij vaste x $\Rightarrow t = \frac{1}{v} d \Rightarrow \widehat{v^{-1}} = \frac{\overline{dt}}{\overline{d^2}} \qquad V(\widehat{v^{-1}}) = \frac{\sigma_t^2}{N\,\overline{d^2}}$

$$\chi^2 = \sum \frac{1}{0,1^2}(t_i - d_i/v)^2$$

(Deze geeft zeker niet dezelfde waarde als in ①, zelfs al waren de waarden dezelfde, ze is wel opnieuw minimaal)

Oef 7.1

7 in 10^6 kg H in 1 jaar, geef 90% conf. int voor λ en voor τ

① geen achtergrond $\quad x = 7 \quad$ ⇝ Tabel 7.1: zoek de 90% confidenties

(Let op, deze tabel werkt dus anders dan die voor Gauss! Daar kijk je naar 2zijdige 90%)

$$\lambda^+_{90\%} = 11,77 \qquad \lambda^-_{90} = 3,89$$

wat betekenen deze? ALS $\lambda = 11,77$ is, heb je maar 10% kans dat je $x \leq 7$ vindt. Dus omgekeerd: als je $x = 7$ vindt, hoe groot zou de echte λ dan moeten zijn opdat je maar 10% kans hebt op $x \leq 7$.

Hoe bepaal je nu de halfwaardetijd? $\quad N_p \sim N_{p,0}\, e^{-t/\tau}$

$$\Rightarrow \frac{1}{2} = e^{-t_{1/2}/\tau} \Rightarrow t_{1/2} = +\ln(2)\,\tau$$

Wat is τ? $\quad \tau = 1/\lambda_p \quad$ (de gemiddelde levensduur v. een proton)

$$\text{en } \lambda_p = \frac{\lambda}{N_p} = \frac{\lambda\, M(g/mol)}{m(g) \cdot 6,022 \cdot 10^{23}(1/mol)} = \frac{\lambda}{10^9 \cdot 6,022 \cdot 10^{23}}$$

$$\lambda^- \rightsquigarrow \lambda_p^- = 0,648 \cdot 10^{-32}\ (1/\text{jaar}) \Rightarrow t^-_{1/2} = \ln(2)/\lambda_p^- = 1,07 \cdot 10^{32} \text{ jaar}$$

$$t^+_{1/2} = 0,355 \cdot 10^{32} \text{ jaar}$$

② Er is een achtergrond van 3 evenementen $\Rightarrow (B+\lambda)^\pm = \begin{cases} 11,77 \\ 3,89 \end{cases}$

$$\Rightarrow \lambda^\pm = \begin{cases} 8,77 \\ 0,89 \end{cases}$$

! Dus niet $x = 4$ zetten !

Je kijkt dus eerst naar de poissonverdeling van de twee samen, dan pas pas je aan voor bijkomende effect

(3) Achtergrond van 81 jaar

Analoog: λ_{90}^- bestaat niet

$\lambda_{90}^+ = 3,77$

Oef 7.2

per 5 in een doos

Aantal defect/doos	0	1	2	3	4	5	
Aantal dozen	31	42	18	6	2	1	$\rightarrow \Sigma = 109$

Bepaal 90% confid. interval

De fouten zijn binomiaal verdeeld. Bekijk apart, niet per doos.

$B(x, n=500, p=?)$ je kan dit (zeer goed) benaderen door een Gaussverdeling met $\mu = np$ en $\sigma = \sqrt{np(1-p)}$

Zoek dus μ_{90}^{+-} en leid hieruit p_{90}^{+-} af

$\rightarrow$ 90% centraal confidentie-interval : $\pm Z = \pm 1,64$

$\mp 1,64 = \dfrac{x - np^\pm}{\sqrt{np^\pm(1-p^\pm)}}$ dit kan je analytisch oplossen, maar je kan ook een schatting voor σ gebruiken:

p^+ is grootst $\Rightarrow x - np^+$ is kleinst

$\hat{p} = \dfrac{x}{n} = 0,218 \rightsquigarrow$ vul in in de noemer

$\Rightarrow p^\pm \approx \dfrac{1}{n}\left[109 \pm 1,64\sqrt{n\hat{p}(1-\hat{p})}\right] = \begin{cases} 0,248 \\ 0,188 \end{cases}$

Om exact te werken (dus zonder de schatting volgens 7.4.4) werk je met (7.5)

$\alpha = (1-0,90) \rightsquigarrow Z_{\alpha/2} = Z_{0,05} = 1,64$ $\quad \pi \rightarrow p \quad \sigma = \frac{1}{n}\sqrt{np(1-p)}$

$\hat{p} = \frac{x}{n} = 0,218$ (opnieuw dus deze schatting nodig) $\quad = \sqrt{p(1-p)/n}$

$$\left[\frac{(1000 \cdot 0,218 + 1,64^2) \pm 1,64\sqrt{1,64^2 + 2000(0,218)(0,782)}}{(1000 + 2 \cdot 1,64^2)}\right] = [0,189 ; 0,250]$$

Oef 7.3 $\quad N = 20$, normaal verdeeld $\quad$ 95% conf. $[17,66 ; 22,34]$ voor μ

1) 90% interval? geg $= [\mu \pm z_{0,025}\frac{\sigma}{\sqrt{20}}] \Rightarrow 22,34 - 17,66 = \frac{2\sigma}{\sqrt{20}} \cdot 1,96$

Dus nu $Z_{0,05} = 1,64 \Rightarrow \frac{2\sigma}{\sqrt{20}} \cdot 1,64 = \frac{(22,34 - 17,66)}{1,96} \cdot 1,64 = 3,92$

Nieuw interval is dus $[\mu - \frac{3,92}{2} ; \mu + \frac{3,92}{2}]$ en $\mu = 20 \Rightarrow [18,04 ; 21,96]$

2) $N_2 = 10N_1 \Rightarrow \sigma_2(\bar{x}) = \frac{\sigma(x)}{10N_1} = \sigma_1(\bar{x})/\sqrt{10} \Rightarrow$ breedte wordt $\frac{2\sigma_1}{\sqrt{10}\sqrt{200}} 1,96 \Rightarrow$ 10 keer smaller

Oef 8.1 $\{3,9; 4,5; 5,5; 6,1\}$ $\mu = 4,9$ Da $x_s = 7,3$

Waarschijnlijk dat deze uit dezelfde steekproef komt?

Zoek de t-waarde van x_s?

σ onbekend : $\hat{\sigma} = \sqrt{\frac{1}{N} \sum_i (x_i - \mu)^2} = 0,86$

$$\Rightarrow t(7,3) = \frac{(7,3 - 4,9)}{0,86} = 2,79\sigma$$

Zoek nu in de tabel de significantie hiervan met $N = 4$

$3,747 > t > 2,776$ ⇒ We kunnen H_0 ($x_s \in$ zelfde steekproef) op het 97,5% niveau (eenzijdig) of 95% (tweezijdig, wat hier eigenlijk de juiste is) verwerpen!
Maar niet op het 99% (tweezijdig).

Er is dus slechts een kans, kleiner dan 5%, dat ze van deze verdeling afkomstig is.

Oef 8.2

1) 50 kaarten, 32 juist → Betekenisvol?

Dit is Bernouilli, maar bij grote N ongeveer Gauss, met

$\mu = np$ en $V = np(1-p)$

$p = 0,5$ $n = 50$ ⇒ $\mu = 25$ $\sigma = \sqrt{25 \cdot 0,5} = 3,54$

$$z(32) = \frac{32-25}{3,54} = 1,98$$

H_0 is dat aantal juist niet groter zal zijn dan 25

⇒ eenzijdig hier!

Zoek in Gauss Tabel eenzijdige significantie:

$P(z \geq 1,98) = 100 - 97,61$ % ⇒ Dus is significant op het $\alpha = 95\%$ niveau (Dan verwerp je H_0), zelfs op 97,5

Maar niet significant op het $\alpha = 99\%$ niveau (verwerp niet)

2) Opdat wél $99\% = \alpha$, zou $z \geq 2,32$

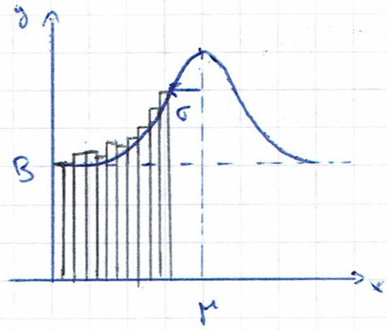
$$\Rightarrow n_j \geq 25 + 2,32 \cdot 3,54 = 33,2$$

Hij zou dus minstens 34 juiste kaarten moeten raden.

Oef 3 $1000 = N$ over 25 Bins ($\sim N$) aan Binwaarde is Gauss gefit en vlakke achtergrond.

Aantal vrijheidsgraden ?

Oorspronkelijk 25, maar er worden 2 parameters gefit voor de Gauss, en 1 voor de achtergrond $\Rightarrow$ 22 vrijheidsgraden



Oef 8.4

$\sigma = 0,2\,K$

10,2	10,4	9,8	10,5	9,9	9,8	10,3	10,1	10,3	9,9

↳ allemaal Gaussisch verdeeld rond T_0

1) H_0 : allemaal metingen van dezelfde temperatuur

μ is onbekend $\rightarrow$ we gebruiken $\bar{x}$ $(N-1)$

Dus zijn er nog 9 vrijheidsgraden over.

$$\chi^2 = \sum_i \frac{(x_i - \bar{x})^2}{\sigma^2} \qquad \bar{x} = 10,12$$

$$= \frac{1}{0,04} \sum_i (x_i - 10,12)^2 = 14,9$$

(Snel? x_i in L_1, $L_2 = (L_1 - 10,12)^2$
Varstats $L_2 \rightarrow \Sigma x \rightarrow \frac{}{0,04}$
Σx^2 als je $L_2 = L_1 - 10,12$)

Nu vind je in de tabel voor χ^2 met $n = 9$: $14,9 > 14,67$
$14,9 < 16,92$

Je zou dus H_0 verwerpen op het $\alpha = 90\%$ niveau, niet op het 95% niveau

(opm: Tabel 7.3 is, zoals het altijd hoort bij χ^2, eenzijdig ☺)

2) Als $\mu = 10,1\,K$ gaat het aantal vrijheidsgraden dus weer naar 10

$$\chi^2 = \frac{1}{0,04} \sum_i (x_i - 10,1)^2 = 15 < 15,99$$ dus je kan H_0 niet verwerpen op het 90%-niveau

Oef 8.5

$n_1 = 11 \quad \bar{x}_1 = 50,9033\,\mu m \quad s_1^2 = 0,1848\,\mu m^2$

$n_2 = 8 \quad \bar{x}_2 = 50,2525\,\mu m \quad s_2^2 = 0,2837\,\mu m^2$

$H_0 : \mu_1 = \mu_2$

2 steekproeven gelijk? σ (en μ) zijn geschat $\Rightarrow$ Student t

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{s\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$ ALS $F = \frac{s_2^2}{s_1^2} < 2,95$ (voor $f_1 = 8$, $f_2 = 11$, $\alpha = 0,05$) Dit is duidelijk het geval, dus is $S_2 = S_2$ met 95% confidentie

$$s^2 = \frac{10 \cdot s_1^2 + 7 \cdot s_2^2}{17} = 0,226$$

$\Rightarrow t = 2,94$ voor 17 vrijheidsgraden, dit is meer dan 2,89

$\Rightarrow$ we moeten H_0 met een confidentie van 99% verwerpen