H5: schatten

5.1 eigenschappen van schatters

def: schatter	= een procedure die toegepast wordt op een steekproef en een numerieke waarde geeft voor een eigenschap of een parameter van de ouderverdeling

5.1.1 consistentie, vooroordelen en efficiëntie

def: consistentie van een schatter	Een schatter is consistent als hij naar de echte waarde toegroeit wanneer het aantal metingen naar oneindig gaat: $\lim_{N\to\infty} \hat{a} = a$ maw: het is altijd mogelijk voor een consistente schatter om een kritiek aantal van metingen n te vinden waarvoor de kans dat de schatter $\hat{a}^n$ meer dan ε van de echte waarde a afwijkt, kleiner is dan elke klein aangegeven waarde η
def: vooroordeel ve schatter	een schatter is onbevooroordeeld als zijn verwachtingswaarde gelijk is aan de echte: $< \hat{a} >= a$ > het vooroordeel ve schatter is gedefinieerd als: $B(\hat{a}) = \mid < \hat{a} > - a \mid.$
def: efficiëntie ve schatter	een schatter is efficiënt als zijn variantie V(â) klein is
def: gemiddelde gekwadrateerde afwijking	voor een schatter â is dit de som van zijn variantie en het kwadraat van zijn vooroordeel: $\Upsilon(\hat{a}) = V(\hat{a}) + [B(\hat{a})]^2$ > helpt bij keuze van beste schatter nl: meest efficiënte schatter is altijd bevooroordeeld onbevooroordeelde schatter is inefficiënt > welke is best?

5.2 fundamentele schatters

5.2.1 het gemiddelde schatten

st: variantie ve schatter	Voor N het aantal meetpunten en σ de standaardafwijking vd ouderdistributie Voor ^μ het gemiddelde van de steekproef > deze geeft een schatting voor het echte gemiddelde: $\hat{\mu} = \bar{x}$ > de variantie van de schatter ^μ is dan gelijk aan: $V(\hat{\mu}) = \frac{\sigma^2}{N}$
verschil in standaardafwijkingen	- *standaardafwijking σ* duidt aan hoe waarschijnlijk het is dat een individuele meting zal afwijken van het gemiddelde μ - *standaardafwijking van het gemiddelde σ/√N* duidt aan in hoeverre onze schatting af kan wijken van het echte gemiddelde

5.2.2 de variantie schatten

Stel dat het gemiddelde μ bekend is

schatter bij gekend gemiddelde	Een consistente, onbevooroordeelde schatter bij een bekend gemiddelde μ is: $\widehat{V(x)} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$

stel dat het gemiddelde μ niet bekend is	
schatter bij onbekend gemiddelde	Neem dezelfde schatter als daarvoor, waar vervang deze door de schatt $\hat{\mu} = \bar{x}$: $$\widehat{V(x)} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$$
> verwachtingswaarde vd schatter	$$\begin{aligned} < \widehat{V(x)} > &= \frac{1}{N}\sum_{i=1}^{N} < (x_i - \bar{x})^2 > \\ &= \frac{1}{N}\sum_{i=1}^{N} < x_i^2 - (\bar{x})^2 > \\ &= \frac{1}{N}\sum_{i=1}^{N}[< x_i^2 > - < (\bar{x})^2 >] \\ &= \frac{N}{N}[< x^2 > - < (\bar{x})^2 >] \\ &=< x^2 > - \underbrace{[< x >^2 - < \bar{x} >^2]}_{=0} - < (\bar{x})^2 > \\ &= V(x) - V(\bar{x}) \end{aligned}$$ deze is bevooroordeeld want: dat $V(\bar{x}) = V(x)/N$ zodat $$< \widehat{V(x)} >= V(x) - \frac{V(x)}{N} = \frac{N-1}{N}V(x) \neq V(x)$$ Maar voor grote N is dit verwaarloosbaar
def: correctie van Bessel	dit is een consistent en onbevooroordeelde schatter van de variantie wanneer het onbekende gemiddelde geschat is: $$\widehat{V(x)} = s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$ > dit is dezelfde als daarnet, enkel gecorrigeerd voor de verwachtingswaarde
de variantie van de schatter van de variantie	
st: variantie vd schatter vd variantie 5.2.2	er geldt: $$V(\widehat{V(x)}) = \frac{(N-1)^2}{N^3} < (x- < x >)^4 > - \frac{(N-1)(N-3)}{N^3} < (x- < x >)^2 >^2$$
standaardafwijking schatten	
schatting standaardafwijking	We kunnen schatten: $$\hat{\sigma} = \sqrt{\widehat{V(x)}}$$ deze blijft consistent en praktisch onbevooroordeeld
st: variantie van schatter van st.afw	De schatter van de standaardafwijking $\hat{\sigma}_x = \sqrt{\widehat{V(x)}}$ heeft als variantie : $$V(\hat{\sigma}) = \frac{1}{4N\sigma^2}\left[< (x- < x >)^4 > - < (x- < x >)^2 >^2\right]$$

5.2.3 de correlatiecoëfficiënt schatten	
schatting van correlatie	De correlatiecoëfficiënt in een steekproef kan gebruikt worden als een schatting voor de correlatie in de ouderdistrubutie: $$\hat{\rho} = \frac{\overline{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y}$$ verwijder de bias via de correctie van Bessel: $$\hat{\rho} = r = \frac{1}{(N-1)} \sum_{i=1}^{N} \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right) = \frac{N}{(N-1)} \frac{\overline{xy} - \bar{x}\bar{y}}{s_x s_y}$$
fout op ρ	voor N groot genoeg kunnen we aantonen dat $$\sigma_\rho = \frac{1-\rho^2}{\sqrt{N-1}}$$ Beter is eigenlijk: $$z = \frac{1}{2} \ln \frac{1+\hat{\rho}}{1-\hat{\rho}}$$
5.3 stratified sampling	
5.3.1 √N verslaan	
stratified sampling	= als een bepaalde bevolking op te delen valt in twee, waarvan je de relatieve verhouding kent en deze types hebben verschillend gemiddelde > neem dan een steekproef met een vast aantal van elk type >> we bekijken eerst wat er gebeurt als we dit niet doen:
variantie bij willekeurige sampling	Zij er een distributie van een mengeling van $P_1(x)$ en $P_2(c)$ Zij respectievelijk de gemiddelden μ_1, μ_2 en verhoudingen f_1,f_2 > het gemiddelde is dan: $$\mu = f_1\mu_1 + f_2\mu_2$$ en de variantie: $$V(x) = \int (f_1 P_1(x) + f_2 P_2(x))\ (x-\mu)^2 dx$$ Dit is iets moeilijker te berekenen dan het eruit ziet. De distributie van type 1 $P_1(x)$, heeft μ_1 als gemiddelde. De variantie van $P_1(x)$ is dus gedefiniëerd als : $$V_1(x) = \int P_1(x)(x-\mu_1)^2 dx$$ Hetzelfde geldt ook voor type 2: $$V_2(x) = \int P_2(x)(x-\mu_2)^2 dx$$
st: variantie van som	De verwachte variantie van het totaal is de som van de varianties PLUS een term afkomstig van mogelijke (binomiale) fluctuaties in samenstelling van de steekproef $$V(x) = f_1 V_1(x) + f_2 V_2(x) + f_1 f_2 (\mu_1 - \mu_2)^2$$
variantie bij stratified sampling	Bij stratified sampling valt de laatste term in de stelling weg > je verslaat de √N regel > vermindert variaties in je steekproef die afkomstig zijn vd verschillende mogelijke aantallen van elk type die je toevallig uitpikt

5.3.2 wat is de beste stratificatie?	
is stratificatie zinvol?	Veronderstel er zijn m_1 en m_2 metingen van resp. type1 en type2 > maak een schatting voor μ_1 en μ_2 met varianties V_1/m_1 en V_2/m_2 > deze worden gecombineerd tot: $$\mu = f_1\mu_1 + f_2\mu_2$$ met variantie: $$V(\hat{\mu}) = f_1^2 \frac{V_1}{m_1} + f_2^2 \frac{V_2}{m_2}$$ Kies m_1, m_2 zodat V minimaal is > hierbij geldt er altijd m_1+m_2 = N een vast getal > we vinden: $$\frac{m_1}{m_2} = \frac{f_1\sqrt{V_1}}{f_2\sqrt{V_2}} = \frac{f_1\sigma_1}{f_2\sigma_2}$$ > - als σ dezelfde zijn, neem zelfde verhouding in streekproef als in bevolking - als σ verschillend, besteed een groter deel aan de steekproef aan de grotere σ
st: variantie bij k types	Als er k verschillende types zijn wordt de variantie gegeven door $$V = \sum_{i=0}^{k} f_i^2 \frac{V_i}{m_i}$$ > kan geminimaliseerd worden via multiplicatoren van Lagrange > neem als grensvoorwaarde $\sum_i^k m_i = N$ constant zodat: $$\frac{\partial N}{\partial m_i} = 1 \qquad \forall i$$ wordt dit $$m_i \propto f_i\sigma_i$$ zoals je al verwachtte.
5.4 likelihood functie	
5.4.1 likelihood	
def: likelihood	= de waarschijnlijkheid om een steekproef $\{x_1,\ldots,x_N\}$ te hebben als een waarschijnlijkheidsverdeling P(X,a) van een toevallige variabele W wordt voorzien: $$\mathcal{L}(a;x_1,x_2,\ldots,x_N) = P(x_1;a)\times P(x_2;a)\times\cdots\times P(x_N;a) = \prod_{i=1}^{N} P(x_i;a)$$ betekenis: totale waarschijnlijkheid (of waarschijnlijkheidsdichtheid) dat een bepaalde groep van gegevens x_i geproduceerd wordt door een mogelijke waarde van de variabele a
st: verwachtingswaarde ve schatter 5.4.1	voor een schatter a is de verwachtingswaarde: $$<\hat{a}> = \int \hat{a}\mathcal{L}\, dX$$ opmerking: dit is een functie van a, dus bijvoorbeeld: $$<\hat{a}^2> = \int \hat{a}^2\mathcal{L}\, dX$$
5.4.2 consistentie en bevooroordeling	
voorwaarde consistentie	In deze notatie wordt de consistentievoorwaarde: $$\lim_{N\to\infty} <\hat{a}-a> = 0$$
corrigeren voor bias	stel dat er een bias is: $$<\hat{a}> = b$$ dan is (â-b) unbiased
bias via consistentie ve schatter	als een schatter consistent is verdwijnt de bias voor N→∞

5.4.3 efficiëntie: minimum variance bound	
efficiëntie van een schatter	De efficiëntie van een schatter is afhankelijk van situatie tot situatie > een schatter is efficiënt als de spreiding vd distributie klein is rond het gemiddelde: $$V(\hat{a}) =< (\hat{a}-a)^2 >=< \hat{a}^2 > - < \hat{a} >^2$$
th: Minimum Variance Bound 5.4.1	Voor een onbevooroordeelde schatter is de MVB: $$V(\hat{a}) \geq \frac{1}{\left\langle \left(\frac{\mathrm{d}\ln\mathcal{L}}{\mathrm{d}a}\right)^2 \right\rangle} \quad (5.18)$$ hetgeen ook kan gescheven worden als: $$V(\hat{a}) \geq \frac{-1}{\left\langle \left(\frac{\mathrm{d}^2\ln\mathcal{L}}{\mathrm{d}a^2}\right) \right\rangle} \quad (5.19)$$ waarbij $\mathcal{L}$ de likelihood functie is, gedefiniëerd in Definitie-5.4.1. Als voor een bepaalde schatter $\hat{a}$ de $V(\hat{a})$ gelijk is aan de MVB, dan zeggen we dat $\hat{a}$ een 'efficiënte' schatter is. Indien niet, dan is de 'efficiëntie' gelijk aan MVB/$V(\hat{a})$. > geeft een limiet aan de nauwkeurigheid van een schatter
st: MVB voor biased schatter	Voor een biased schatter met bias b wordt de teller in MVB vervangen door $(1+db/da)^2$
def: informatie I(a)	= de grootheid in de noemer van de MVB: $$I(a) = \left\langle \left(\frac{d\ln\mathcal{L}}{da}\right)^2 \right\rangle$$
I(a) herschreven	we kunnen I(a) ook voorstellen als: $$I(a) = -N\int \frac{d^2\ln P}{da^2} P dx = N\int \left(\frac{d\ln P}{da}\right)^2 P dx$$
5.4.4 Gaussische verdeling	
MVB van een Gauss	Voor een Gaussische verdeling is het gemiddelde van een steekproef een efficiënt schatter van μ. De waarschijnlijkheid voor een bepaalde x_i is $$P(x_i;\mu) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$ hetgeen leidt tot een totale 'waarschijnlijkheid', of likelihood (en we nemen meteen de logaritme van de likelihood) $$\ln\mathcal{L} = -\sum_i^N \frac{(x_i-\mu)^2}{2\sigma^2} - N\ln(\sqrt{2\pi}\sigma)$$ Dit tweemaal afleiden naar μ geeft $$\frac{d^2\ln\mathcal{L}}{d\mu^2} = -\frac{N}{\sigma^2}$$ In deze uitdrukking komt x niet voor, zodat ze onmiddellijk gelijk is aan haar verwachtingswaarde. Inverteren en van teken veranderen geeft $$\mathrm{MVB} = \frac{\sigma^2}{N}$$ hetgeen eveneens de variantie van de schatter is.