

De drie mooiste modellen deel 1

Medische statistiek Basisbegrippen

Hans Burgerhof
h.burgerhof@home.nl

Doel en middelen

- Medische statistiek proberen inzichtelijk te maken aan de hand van artikelen uit de medische literatuur, met behulp van realistische voorbeelden en concrete datasets.
- Er wordt gebruik gemaakt van (powerpoint) presentaties, een syllabus, SPSS en opdrachten aan de hand van geselecteerde artikelen

Programma

- **Donderdag**
- Blok 1: definitie van statistiek; beschrijvende statistiek en kansrekening
- Blok 2: vervolg kansrekening; inferentiële statistiek
- Blok 3: diverse statistische toetsen en diagnostische testen
- Blok 4: gastspreker
- **Vrijdag**
- Blok 5: enkelvoudige lineaire regressie
- Blok 6: meervoudige lineaire regressie
- Blok 7: logistische regressie
- Blok 8: survival analyse

Opbouw syllabus

- Deel 1: continue responsievariabele
 - diverse statistische toetsen, lineaire regressie
- Deel 2: dichotome responsievariabele
 - diverse statistische toetsen, logistische regressie
- Deel 3: survival analyse

Statistiek is

De wetenschap die zich bezig houdt met

- Het opzetten van onderzoek (methodologie)
- Het verzamelen van data
- Het ordenen, samenvatten en presenteren van deze data
- Het analyseren van deze data
- Het trekken van (statistische) conclusies

kansrekening

Beschrijvende statistiek

Doelen:

- beschrijven populatie
- gevoel krijgen voor je data
- controleren van de ingevoerde gegevens

Maten van ligging

- modus
- mediaan
- gemiddelde

Maten van spreiding

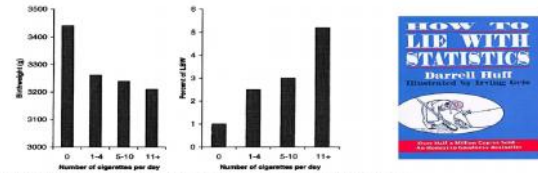
- spreidingsbreedte
- kwartiel afstand
- standaarddeviatie

Table 1 from Phung et al. (n = 3242)

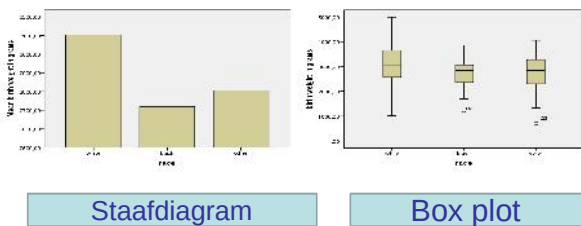
Page 31.6% smoked: $n_s = 1024$

	Smokers	Non-smokers	Prevalence of smoking*
Mother's age			
< 20	190(1.1)	85(4.0)	51.4 ± 3.8
20-29	61(0.4)	1150(33.2)	34.7 ± 1.1
30-39	278(7.7)	875(40.6)	34.2 ± 1.3
40+	161(6.3)	41(1.9)	28.1 ± 6.0
Marital status			
Married	429(42.7)	1829(76.1)	20.8 ± 0.9
Single	236(24.9)	193(9.0)	36.4 ± 2.4
Living with partners	329(32.7)	318(14.9)	50.1 ± 2.0
Race			
Caucasian	939(96.5)	168(75.0)	36.4 ± 0.9
Asian	150(5)	143(6.5)	3.4 ± 1.5
Middle Eastern	11(0.1)	31(2.3)	21.5 ± 5.1
Pacific Islanders	11(0.1)	47(3.1)	16.5 ± 6.1
African and others	167(6.5)	266(13.1)	18.9 ± 1.2
Parity			
1	366(36.2)	640(39.9)	30.3 ± 1.3
2-3	300(30.3)	1091(51.8)	31.8 ± 1.2
≥ 4	136(3.5)	176(8.3)	43.6 ± 2.8
Maternal care			
Antenatal care	451(69.3)	743(55.5)	37.8 ± 1.4
Obstetric care	94(14.6)	427(31.9)	18.2 ± 1.7
Shared care	91(14.0)	158(11.6)	37.6 ± 3.1
Others	114(2.2)	14(1.1)	50.0 ± 9.4
Education			

Figure 1, page 239



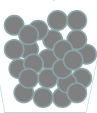
Twee grafieken



P(bal = rood)?

Kansrekening:
We kennen de inhoud van
de vaas

?



Statistiek:
We weten de inhoud van de vaas niet

Wat zegt de steekproef over de
inhoud van de vaas?

Wat zegt de proportie zieken in de
steekproef over de prevalentie in de
populatie?

Om inferentiële
statistiek te kunnen
begrijpen, hebben
we kennis van
kansrekening nodig

De klassieke kansdefinitie

- Pierre-Simon Laplace (1774)

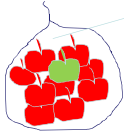
$$P(\text{gebeurtenis } A) = \frac{\text{mogelijk aantal gunstige uitkomsten voor } A}{\text{totaal aantal mogelijke uitkomsten}}$$

P kort voor "probability"

$$P(5) = 1/6$$

$$P(\text{even}) : 3/6 = 0,5$$

Een vrouw is zwanger (een baby).
Wat is de kans dat de baby een jongetje is?

 $P(\text{appel} = \text{groen})?$

$P(\text{appel} = \text{groen}) = 1 / 10 = 0,1$
(want er zijn tien appels)

$P(\text{appel} = \text{groen}) = 1 / 2 = 0,5$
(want er zijn maar twee kleuren) ?

De klassieke kansdefinitie

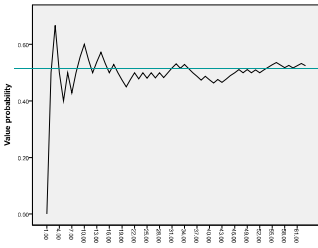
 $P(\text{gebeurtenis } A) = \frac{\text{mogelijk aantal gunstige uitkomsten voor } A}{\text{totaal aantal mogelijke uitkomsten}}$

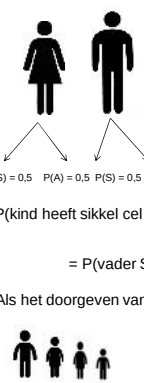
Als alle mogelijke uitkomsten even waarschijnlijk zijn!

 $P(\text{jongen}) = \frac{1}{2}?$

Empirische kansdefinitie

$$P(\text{gebeurtenis } A) = \lim_{n \rightarrow \infty} \frac{\text{aantal uitkomsten } A}{n}$$

 $P(\text{jongen}) \approx 0,52$ (in Nederland)

 Sikkel cel drager: heterozygoot AS

$P(S) = 0,5 \quad P(A) = 0,5 \quad P(S) = 0,5 \quad P(A) = 0,5$

$P(\text{kind heeft sikkel cel ziekte}) = P(\text{vader en moeder geven beiden } S \text{ door})$

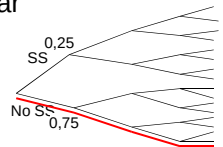
$= P(\text{vader } S) \cdot P(\text{moeder } S) = 0,5 \cdot 0,5 = 0,25$

Als het doorgeven van genen onafhankelijke gebeurtenissen zijn

Deze ouders hebben vier kinderen. Wat is de kans dat **precies één** van hen de ziekte heeft?

De kans op d "successen" in 4 observaties, alle met kans 0,25 op "succes"

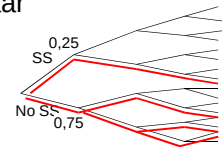
- We nemen aan dat alle kinderen dezelfde kans op de ziekte hebben, onafhankelijk van elkaar



- $P(0 \text{ zieke kinderen}) = 0,75^4$

De kans op d "successen" in 4 observaties, alle met kans 0,25 op "succes"

- We nemen aan dat alle kinderen dezelfde kans op de ziekte hebben, onafhankelijk van elkaar



- $P(1 \text{ kind met de ziekte}) = 4 \cdot 0,25 \cdot 0,75^3$

De Binomiale verdeling met n onafhankelijke observaties, alle met succeskans π

We have a Binomial experiment if we have a binary outcome (success or failure) with probability of success π and n independent observations. The number of successes X has a Binomial distribution: $X \sim B(n, \pi)$
The probability of exactly d successes in n observations:

$$P(X = d) = \binom{n}{d} p^d (1-p)^{n-d}$$

In which $\binom{n}{d} = \frac{n!}{(n-d)!d!}$ with n factorial ($n!$) = $n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$

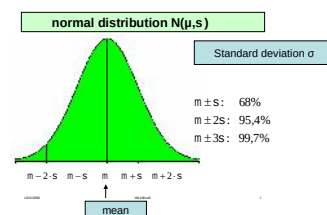
Voorbeelden met een discrete uitkomst

- Worp met een eerlijke dobbelsteen: $P(\text{aantal ogen} = 4)$?
- De kans op zwangerschap bij IVF = 0,3. Hoe groot is de kans dat twee vrouwen die beiden onafhankelijk van elkaar IVF ondergaan beiden zwanger worden?
- Hoe groot is de kans dat van tien vrouwen die IVF ondergaan er minstens 4 zwanger worden?

Kansrekening (2)

- Kansen in een continue verdeling
- Gegeven een populatie met een normaal verdeelde bloeddruk met $m = 120$ en $s = 10$ mm Hg
- Hoe groot is de kans dat bij een willekeurig gekozen persoon uit deze populatie de bloeddruk hoger is dan 140?

Probability distributions

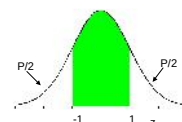


Uitwerking

- $P(X > 140) = P(Z > (140 - 120)/10) = P(Z > 2) = 0,023$
- Het vertalen van het kansprobleem naar de standaardnormale verdeling heet standaardiseren. Je trekt het gemiddelde van de gezochte grens (140) af en je deelt vervolgens door de standaarddeviatie. Je krijgt dan het gezochte verschil in aantallen standaarddeviaties.

Normal distribution, two-sided P

z	P	z	P
0.0	1.00	1.6	0.11
0.1	0.92	1.645	0.100
0.2	0.84	1.7	0.089
0.3	0.76	1.8	0.072
0.4	0.69	1.9	0.057
0.5	0.62	1.96	0.050
0.6	0.55	2.0	0.046
0.674	0.500	2.1	0.036
0.7	0.48	2.2	0.028
0.8	0.42	2.3	0.021
0.9	0.37	2.4	0.016
1.0	0.32	2.5	0.012
1.1	0.27	2.576	0.010
1.2	0.23	3.0	0.0027
1.3	0.19	3.291	0.0010
1.4	0.16		
1.5	0.13		



Kansrekening (3)

- Er worden 25 personen uit de normaal verdeelde populatie getrokken
- Wat is de kans dat het gemiddelde hoger is dan 140 mm Hg?
- Centrale limietstelling
- "Standaard error" = $sd/\sqrt{n}$

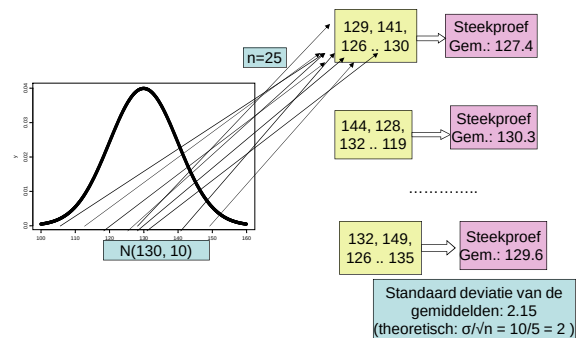
Centrale limiet stelling

1. De verwachte waarde van het gemiddelde van steekproefgemiddelden is het populatiegemiddelde
2. De standaarddeviatie van de steekproefgemiddelden is $\frac{s}{\sqrt{n}}$ waarin σ de sd. In de populatie is
3. de verdeling van de gemiddelden is normaal als de oorspronkelijke verdeling normaal verdeeld was
4. Maar ook als de oorspronkelijke verdeling niet normaal is, is de verdeling van de gemiddelden bij benadering normaal, voor grote n

Illustratie

- We willen de gemiddelde bloeddruk schatten van een populatie
- Random simulatie uit $N(130, 10)$
129, 141, 126, ... ,130.
steekproefgemiddelde: 127.4 , sd = 9.98
- een tweede simulatie: gem.: 130.3, sd = 7.75
- En nog drie gem.: 130.5 , 125.5, 129.6
- Sd van deze vijf gemiddelden: 2.15

Wat is er gebeurd?



$P(X_{\text{gemiddeld}} > 140)?$

- Als de 25 waarnemingen uit een normale verdeling komen met een gemiddelde van 120 en een sd van 10, geldt: $P(X_{\text{gem}} > 140)$
 $= P(Z > (140 - 120)/(10/\sqrt{25})) = P(Z > 20/2)$
 $= P(Z > 10) < 0,001$

Verklarende statistiek

- Schatten en toetsen
 - een onbekend gemiddelde μ (one-sample t-toets)
 - verschil van twee gemiddelden (t-toets voor onafhankelijke groepen, Mann-Whitney, gepaarde t-toets)
 - verschil van twee proporties (chi-kwadraat toets)
 - ...

Schatten van een onbekend gemiddelde μ bij een normaal verdeelde variabele

- Neem een aselechte steekproef van n personen
- Schat μ met behulp van het steekproefgemiddelde
- Kies een onbetrouwbaarheid α
- Bereken het $100(1-\alpha)\%$ betrouwbaarheidsinterval
- Wat zegt dat BI? Hoe doe je dat?

Hoe betrouwbaar is die schatting?

- Een betrouwbaarheidsinterval geeft een gebied waarin met % betrouwbaarheid de onbekende populatieparameter ligt
- Hiervoor zijn nodig
 - een gekozen betrouwbaarheid (90%, 95%, 99%) en de bijbehorende z-waarde
 - de standaarddeviatie van de schatter ("standard error")

95% BI voor μ

$$\bar{X} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

- Bijvoorbeeld: onbekende tlc in een gespecificeerde populatie
- steekproefgemiddelde = 6,09 liter
- $n = 32$ steekproef sd = 1,62
- $6,09 \pm 1,96 \cdot 0,29 \approx [5,50 ; 6,67]$

Z-waarde; alleen correct als de populatie Sd bekend is. Anders gebruik t-waarde met $df = n-1$

Toetsen

- Bij toetsen stelt men een nulhypothese (H_0) die het tegendeel van de onderzoeksvraag weerspiegelt, en een alternatieve hypothese. Men kiest een significantieniveau α en berekent de P-waarde. Dit is de kans om de gevonden steekproefwaarden te vinden (of nog extremer) als H_0 waar is. Als $P < \alpha$ verworpt men de nulhypothese.

Verschillen bij onafhankelijke groepen

	Y nominaal	Y ordinaal	Y normaal verdeeld
één groep	Binomiaal, Chi-kwadraat	Tekentoets, Wilcoxon signed rank	T-toets
Twee groepen	Fisher's exacte toets, Chi-kwadraat	Mann-Whitney (Wilcoxon ranksum)	T-toets voor onafhankelijke groepen
Meer dan twee groepen	Chi-kwadraat	Kruskal-Wallis	(one-way) ANOVA

Voorbeeld

- Als je verwacht dat de tlc in een specifieke populatie lager is dan "normaal" (normaal is bv. 7 liter), is de $H_0: \mu = 7$.
- Je verzamelt data en berekent het steekproefgemiddelde (X_{gem}) en de standaarddeviatie van dit gemiddelde (de "standard error")
- Vervolgens bereken je de kans om dit gemiddelde te vinden (of extremer) als de H_0 waar is. Dit is je P-waarde

Eén steekproef t-toets

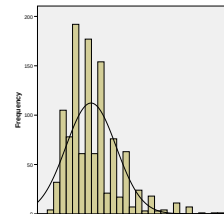
One-Sample Statistics				
	N	Mean	Std. Deviation	Std. Error Mean
total lung capacity (l)	32	6,0878	1,62406	,28710

One-Sample Test					
Test Value = 7					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower Upper
total lung capacity (l)	-3,177	31	,003	-.91218	-1,49777 -3267

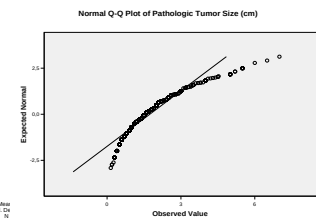
P-waarde
Conclusie?

T-toets: Normaal verdeeld?

Histogram

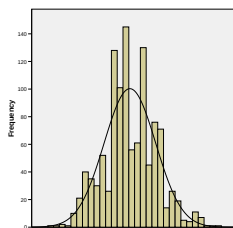


Q-Q Plot

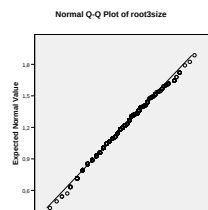


T-toets: Normaal verdeeld?

Histogram



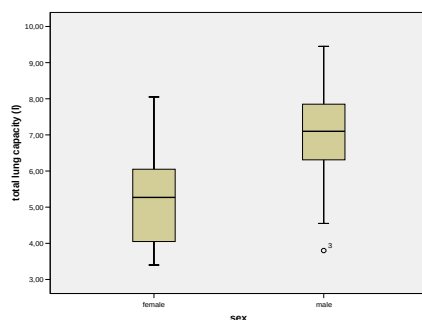
Q-Q Plot



Verschillen tussen twee onafhankelijke groepen

- We willen toetsen of mannen en vrouwen dezelfde gemiddelde tlc hebben
- Nulhypothese?
- Alternatief?
- Welke toets?
- Zijn er voorwaarden voor de toets?
- Significantieniveau?

Boxplot

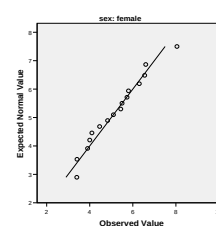


TLC

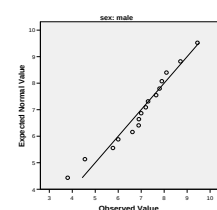
Vrouwen

mannen

Normal Q-Q Plot of total lung capacity (l)



Normal Q-Q Plot of total lung capacity (l)



Toetsen van twee onafhankelijke gemiddelden

- $H_0: \mu_1 = \mu_2$ of $H_0: \mu_1 - \mu_2 = 0$ of $\delta = 0$
- $H_1: \mu_1 \neq \mu_2$ of $H_1: \delta \neq 0$
- Kies α (meestal 0,05)
- Verzamel je gegevens (per groep n)
- Bereken de steekproef gemiddelden en de sd's
- Trek je conclusie (verwerp H_0 of accepteer H_0)
- **NB:** statistische significantie is niet hetzelfde als klinische relevantie!

T-toets voor twee onafhankelijke groepen ($H_0: \mu_1 = \mu_2$)

Group Statistics				
sex	N	Mean	Std. Deviation	Std. Error Mean
total lung capacity (l)				
female	16	5,1981	1,30082	,32521
male	16	6,9775	1,43882	,35970

Independent Samples Test									
Levene's Test for Equality of Variances					t-test for Equality of Means				
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
total lung capacity (l)	.001	.976	-3,669	30	.001	-1,77938	.48492	-2,76971	-.78904
								Lower	Upper
Equal variances assumed									
Equal variances not assumed									

Data moeten normaal verdeeld zijn

Data niet normaal verdeeld?

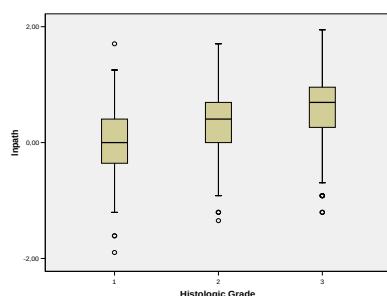
- Helpt het om de data te transformeren?
- Als de getransformeerde data wel normaal verdeeld zijn, gebruik een t-toets op de getransformeerde data en vertaal je resultaten naar de oorspronkelijk schaal
- Zo nee, gebruik een parametervrije toets

Welke? Hoe werkt die ook al weer?

Meer dan twee groepen, normaal verdeelde Y

- Alle groepen paarsgewijs vergelijken?
- Kanskapitalisatie!
- $H_0: \mu_1 = \mu_2 = \dots = \mu_k$
- One-way ANOVA
- Voorwaarden
 - Onderling onafhankelijke waarnemingen
 - Normaal verdeelde responsvariabele
 - Gelijke varianties in de groepen

Ln(tumorgrootte) naar histologische gradering



Beschrijvende statistiek en de voorwaarde van gelijke varianties

Descriptives								
Inpath	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	77	-.0034	.64241	.07321	-.1492	.1424	-1.90	1.70
2	484	.3527	.55995	.02545	.3027	.4027	-1.35	1.70
3	313	.6776	.50858	.02307	.6265	.7287	-1.20	1.94
Total	874	.3998	.59951	.02028	.3600	.4396	-1.90	1.95

Gelijke varianties?

$$H_0: S_1^2 = S_2^2 = S_3^2$$

Test of Homogeneity of Variances

Inpath	Levene Statistic	df1	df2	Sig.
	.147	2	871	.863

De ANOVA (ANalysis Of VAriance)

ANOVA

Inpath	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	22.874	2	11.437	34.245	.000
Within Groups	290.889	871	,334		
Total	313.762	873			

Wat gebeurt er precies?

Wat is de conclusie?

Post-hoc tests Bonferroni

Multiple Comparisons

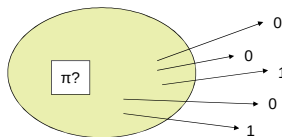
Dependent Variable: Inpath
Bonferroni

(I) Histologic Grade	(J) Histologic Grade	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-.35607	.07090	.000	-.5261	-.1860
	3	-.57539	.07351	.000	-.7517	-.3991
2	1	.35607	.07090	.000	.1860	.5261
	3	-.21932	.04192	.000	-.3199	-.1188
3	1	.57539	.07351	.000	.3991	.7517
	2	.21932	.04192	.000	.1188	.3199

*. The mean difference is significant at the .05 level.

Responsievariabele “laag geboortegewicht”

- Laag geboortegewicht = 0 (geb.gew.>2500 gram)
- Laag geboortegewicht = 1 (geb.gew.<2500 gram)
- Hoe vaak komt dit voor bij voldragen zwangerschappen?



Steekproefproportie p wordt gebruikt als puntschatting voor de onbekende π

Betrouwbaarheidsinterval

- Voor een gemiddelde: $\bar{X} \pm z_{1-\alpha/2} \frac{s}{\sqrt{n}}$
- Wat is de sd van een proportie?
- Terug naar de kansrekening: de binomiale verdeling: als $X \sim B(n, p)$ geldt

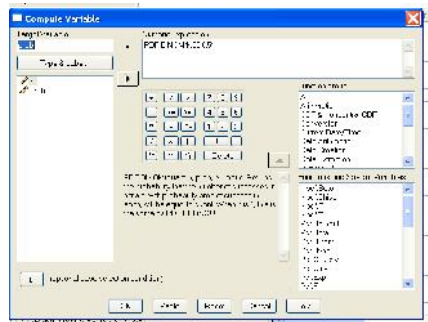
$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$E(X) = np$$

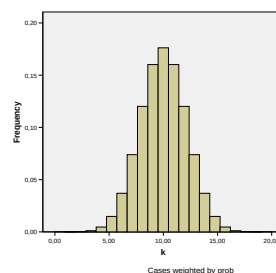
$$Sd(X) = \sqrt{np(1-p)}$$

X = het aantal “successen”

Kansverdelingen in SPSS (Transform – Compute – Function group - PDF)



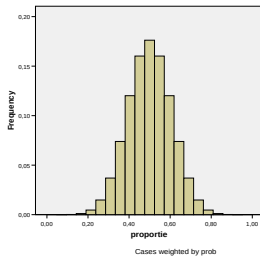
Binomiale kansverdeling n=20, p = 0,5



$$E(X) = np = 20 \cdot 0,5 = 10$$

$$Sd(X) = \sqrt{np(1-p)} = \sqrt{(20 \cdot 0,5 \cdot 0,5)} = 2,24$$

Kansverdeling van een proportie



$$\begin{aligned} \text{Proportie} &= X/n \\ E(\text{proportie}) &= E(X)/n = p \\ \text{Sd}(\text{proportie}) &= \text{sd}(X)/n = \\ &= \sqrt{\frac{np(1-p)}{n}} = \sqrt{\frac{p(1-p)}{n}} \end{aligned}$$

BI voor een proportie (normale benadering)

- 95 % BI: $p \pm 1,96 * \sqrt{\frac{p(1-p)}{n}}$

Voorbeeld:

- We willen de proportie vrouwen schatten die tijdens de zwangerschap roken
- In een steekproef van 189 vrouwen blijken er 74 tijdens de zwangerschap gerookt te hebben
- Puntschatting $p = 74/189 = 0,39$
- Schatting $\text{sd} = \sqrt{\frac{0,39 * 0,61}{189}} = 0,035$

95 % BI:
 $[0,39 - 1,96 * 0,035 ; 0,39 + 1,96 * 0,035]$
 $= [0,32 ; 0,46]$

Waarom zijn de SE's zo verschillend?

Table 1. Characteristics of study subjects

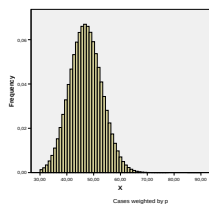
	Smokers	Non-smokers	Prevalence of smoking*
Maternal age			
<20	199(0.1)	85(4.0)	51.4 ± 3.8
20-29	613(61.4)	1136(53.5)	54.7 ± 1.1
30-39	278(27.6)	875(40.6)	28.2 ± 1.3
40+	16(1.6)	41(1.9)	26.1 ± 6.0
Marital status			
Married	428(42.7)	1628(76.1)	20.8 ± 0.9
Single	255(24.9)	103(9.0)	56.4 ± 2.4
Living with partner	725(72.3)	316(4.9)	80.3 ± 2.0
Race			
Caucasian	838(80.3)	1648(73.0)	36.4 ± 0.9
Asian	13(1.3)	14(0.6)	8.4 ± 1.8
Middle Eastern	11(1.1)	51(2.3)	21.5 ± 5.1
Pacific Islanders	11(1.1)	67(3.1)	38.3 ± 4.2
African and others	167(16.3)	286(13.1)	18.9 ± 1.2
Parity			
1	346(36.2)	848(39.0)	16.3 ± 1.3
2-3	598(58.3)	1891(86.8)	51.2 ± 1.2
≥4	134(13.5)	370(3.2)	43.4 ± 2.8
Maternal care			
Antenatal care	431(49.3)	743(55.3)	37.8 ± 1.4
Obstetrical	951(4.4)	427(1.9)	18.2 ± 1.7
Shared care	911(4.6)	159(1.6)	37.0 ± 3.1
Others	118(1.2)	18(1.6)	38.0 ± 9.8

Toetsen

- De overheid wil een campagne starten om het roken tijdens de zwangerschap tegen te gaan, als nu minstens 25 % van de zwangere vrouwen rookt.
- De nulhypothese $\pi = 0,25$ wordt getoetst tegen het alternatief $\pi > 0,25$.
- Steekproef: 74 van de 189 vrouwen rookten tijdens de zwangerschap
- $P(X \geq 74)$ als $X \sim B(189, 0,25)$?

De kansverdeling onder H_0

- Als H_0 waar is, geldt voor het aantal "successen" (rokende zwangeren) X dat $X \sim B(189, 0,25)$
- De P-waarde is dan $P(X \geq 74)$



$$P(X \geq 74) = P(X = 74) + P(X = 75) + \dots + P(X = 189)$$

$$P(X = k) = \binom{189}{k} 0,25^k 0,75^{189-k}$$

Exact antwoord

- Tabellen
- (Grafische) rekenmachine
- SPSS: Transform – Compute –
 $p = 1 - \text{CDF.BINOM}(73, 189, 0,25)$

$$1,266 \dots * E-005 \approx 0,0000127$$

In SPSS

- Analyze – Nonparametric Tests – Binomial

Binomial Test					
	Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (1-tailed)
smoking status during pregnancy	yes	74	,39	,25	,006
	no	115	,61		
	Total	189	1,00		

a. Based on Z Approximation.



Verskil in proporties “lichtgewicht” tussen rokende en niet-rokende moeders

smoking status during pregnancy * low birth weight Crosstabulation

		low birth weight		Total
		no (bwt >= 2500 g)	yes (bwt < 2500)	
smoking status during pregnancy	no	Count 86 % within smoking status during pregnancy 74,8%	Count 29 % within smoking status during pregnancy 25,2%	115
	yes	Count 44 % within smoking status during pregnancy 59,5%	Count 30 % within smoking status during pregnancy 40,5%	74
Total		Count 130 % within smoking status during pregnancy 68,8%	Count 59 % within smoking status during pregnancy 31,2%	189

Puntschatting voor V
0,405 – 0,252 = 0,15

Twee onafhankelijke groepen

- Schatten van een verschil in proporties
- $V = p_1 - p_2$
- $SE(V) = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$
- 95 % BI: $V \pm 1,96 * SE(V)$

$$[0,15 - 1,96 * 0,070 ; 0,15 + 1,96 * 0,070] = [0,01 ; 0,29]$$

Toetsen van $H_0: \pi_1 = \pi_2$

- 1. Chi-kwadraat toets voor onafhankelijkheid
- 2. Fisher's exacte toets (voor de twee-bij-twee tabel)

Chi-kwadraat toets

- In plaats van de nulhypothese te formuleren als “proportie 1 = proportie 2” kun je ook zeggen: de variabelen roken en “laag geboortegewicht” zijn onafhankelijk
- Vervolgens kijken we of de geobserveerde aantallen in de kruistabel overeen komen met de verwachte aantallen als H_0 waar is

Nogmaals de kruistabel

smoking status during pregnancy * low birth weight Crosstabulation

		low birth weight		Total
		no (bwt >= 2500 g)	yes (bwt < 2500)	
smoking status during pregnancy	no	Count 86 % within smoking status during pregnancy 74,8%	Count 29 % within smoking status during pregnancy 25,2%	115
	yes	Count 44 % within smoking status during pregnancy 59,5%	Count 30 % within smoking status during pregnancy 40,5%	74
Total		Count 130 % within smoking status during pregnancy 68,8%	Count 59 % within smoking status during pregnancy 31,2%	189

Twee gebeurtenissen A en B zijn onafhankelijk ⇔
 $P(A \cap B) = P(A) * P(B)$

Toelichting

	A		
B	O		N(b)
	N(a)		N

Verwacht indien onafhankelijk
 $E = P(A) \cdot P(B) \cdot N = [N(a) \cdot N(b)] / N$

De Chi-kwadraat waarde

smoking status during pregnancy * low birth weight Crosstabulation

		low birth weight		
		no (bwt >= 2500 g)	yes (bwt < 2500 g)	Total
smoking status during pregnancy	no	Count 86	29	115
		Expected Count 79,1	35,9	115,0
yes	Count 44	30	74	115
		Expected Count 50,9	23,1	74,0
Total	Count 130	59	189	219
		Expected Count 130,0	59,0	189,0

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = \frac{(86 - 79,1)^2}{79,1} + \frac{(29 - 35,9)^2}{35,9} + \frac{(44 - 50,9)^2}{50,9} + \frac{(30 - 23,1)^2}{23,1} = 4,921$$

Dit wordt vergeleken met een Chi-kwadraat verdeling met 1 df
 Dat geeft een P-waarde. In dit geval 0,026

Chi-kwadraat in SPSS Fisher's exacte toets

Chi-kwadraat is een benadering
 Fisher's toets is exact!

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4,921 ^a	1	,026		
Continuity Correction ^b	4,236	1	,040		
Likelihood Ratio	4,867	1	,027		
Fisher's Exact Test				,036	,020
Linear-by-Linear Association	4,898	1	,027		
N of Valid Cases	189				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 23,10.

Toets voor twee proporties gaf een Z = 2,18
 met een P-waarde van 0,0146 (éénzijdig)

Meer dan twee onafhankelijke groepen

low birth weight * race Crosstabulation

		race			
		white	black	other	Total
low birth weight	no (bwt >= 2500 g)	Count 73	15	42	130
		% within race 76,0%	57,7%	62,7%	68,8%
yes (bwt < 2500 g)	Count 23	11	25	59	115
		% within race 24,0%	42,3%	37,3%	31,2%
Total	Count 96	26	67	189	219
		% within race 100,0%	100,0%	100,0%	100,0%

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Aantal vrijheidsgraden:
 (# rijen - 1) * (# kolommen - 1)

Chi-kwadraat in SPSS

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5,005 ^a	2	,082
Likelihood Ratio	5,010	2	,082
Linear-by-Linear Association	3,570	1	,059
N of Valid Cases	189		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8,12.

Te gebruiken om een trend
 te toetsen bij een ordinale
 groepsindeling
 (in dit geval zinloos)

Tabel 1 van Smit et al.

TABLE 1. Characteristics of children and their mothers			
Characteristic	n	%	P-value
Gender			
Male	130	60,3	0,270
Female	89	39,7	
Maternal age at delivery			
Mean (SD)	31,1 (5,1)		
Range	18-45		
Maternal education			
Mean (SD)	12,1 (2,1)		
Range	9-16		
Maternal employment			
Mean (SD)	1,1 (1,1)		
Range	0-2		
Maternal smoking			
Mean (SD)	1,1 (1,1)		
Range	0-2		
Maternal alcohol use			
Mean (SD)	1,1 (1,1)		
Range	0-2		
Maternal depression			
Mean (SD)	1,1 (1,1)		
Range	0-2		

Welke toetsen zijn hier uitgevoerd?
 Wat is de conclusie?
 Kunnen we deze toetsen
 reproduceren?