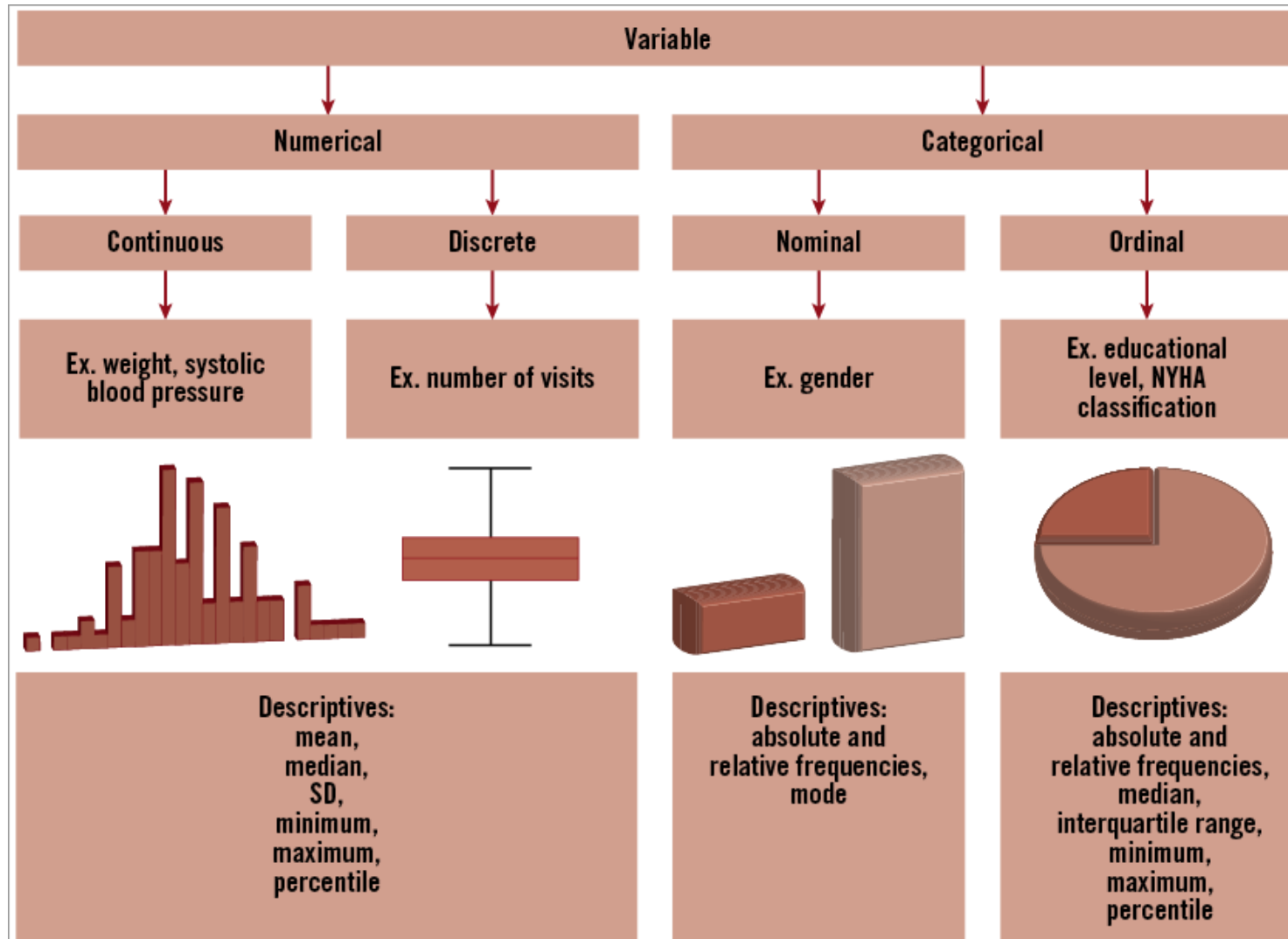


Beschrijvende statistiek



Beschrijvende en toetsende statistiek

- **Beschrijvend** Samenvatting van gegevens in de steekproef van onderzochte personen (gemiddelde, de standaarddeviatie, tabel, grafiek)

- **Toetsend** Uitspraken over de onbekende en ongemeten bredere populatie van individuen (betrouwbaarheidsinterval, statistische tests, p-waarde)

Hypothese toetsen

- Vergelijking is het meest voorkomende elementaire principe in medisch onderzoek. Een verklaring over de waarheid wordt vergeleken met een referentie (de nul)
- **H0: Nul-hypothese**
→ cholesterol is vergelijkbaar tussen mannen en vrouwen
- **H1: Alternatieve hypothese**
→ cholesterol is verschillend tussen mannen en vrouwen
- Kans op het verkrijgen van de waargenomen gegevens in de steekproef gegeven dat de nulhypothese waar is → **P-value**

Significantie en klinische relevantie

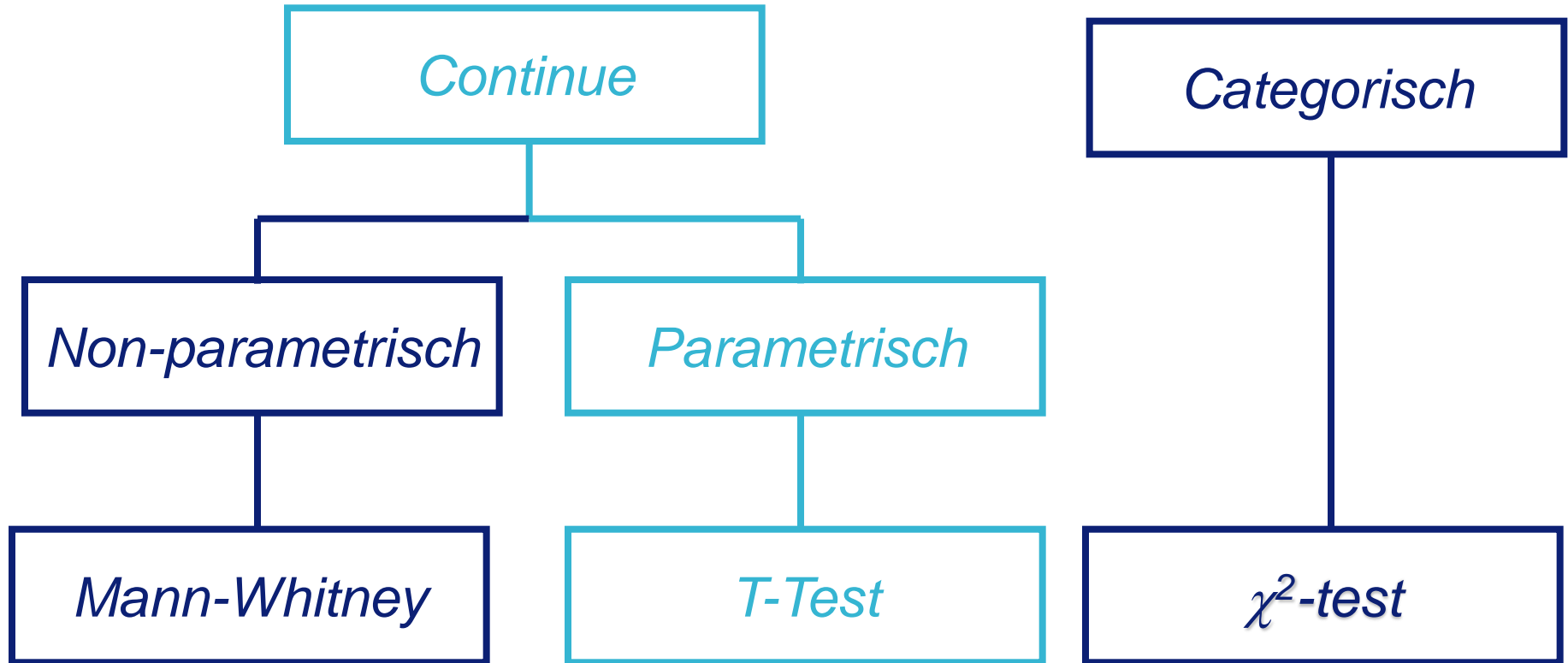
- Als n groot genoeg is, kan zelfs een klein verschil van 0,1 mm Hg statistisch significant tussen de twee groepen.
-> 0.1 mmHg is niet relevant vanuit klinisch oogpunt
- Als n te klein is, kan ook een gemiddelde van 150 mmHg statistisch niet significant verschillend van 130 mm Hg
-> 20 mmHg is relevant vanuit klinisch oogpunt

Conclusies en Realiteit

	Realiteit	H0 is waar	H0 niet waar
Conclusie			
H0 accepteren			TYPE II FOUT β
H0 verwerpen		TYPE I FOUT α	

Vergelijken twee gemiddelden

Vergelijking twee gemiddelden



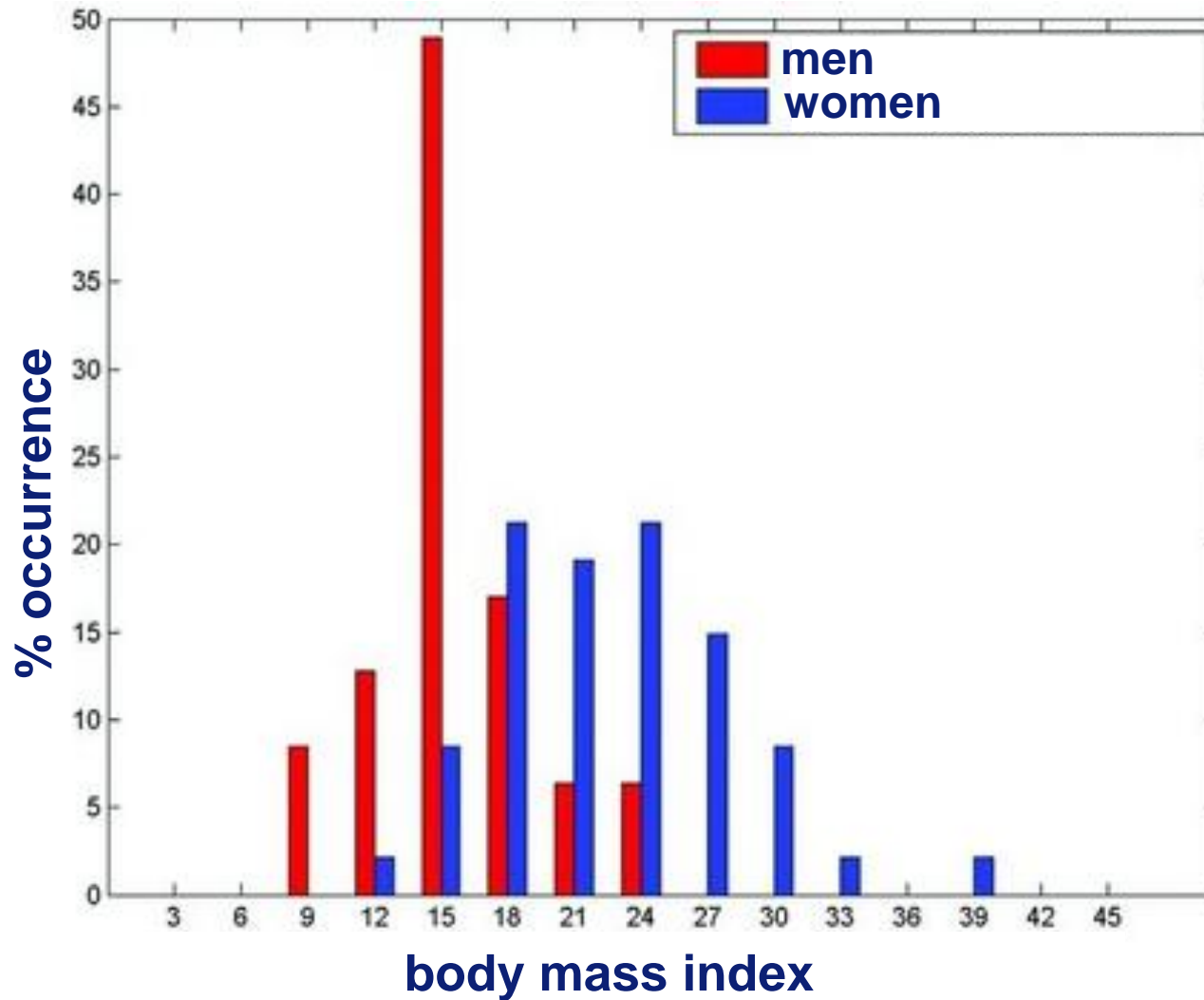
Independent-samples t-test

- Onderzoeksvraag: Is er een significant verschil in de variabele tussen de groepen?
- Twee variabelen:
 - één categorische, onafhankelijke variable (groep)
 - één continue, afhankelijke variabele
- $H_0: \mu_1 = \mu_0$
 $H_1: \mu_1 \neq \mu_0$
- Test vertelt of er een statistisch significant verschil in de gemiddelden voor de twee groepen
- Aanname: gelijke varianties

Independent-samples t-test - voorbeeld

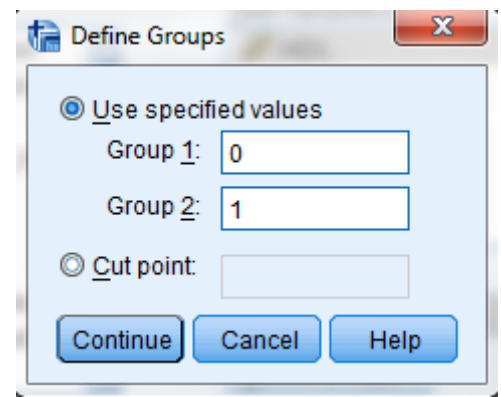
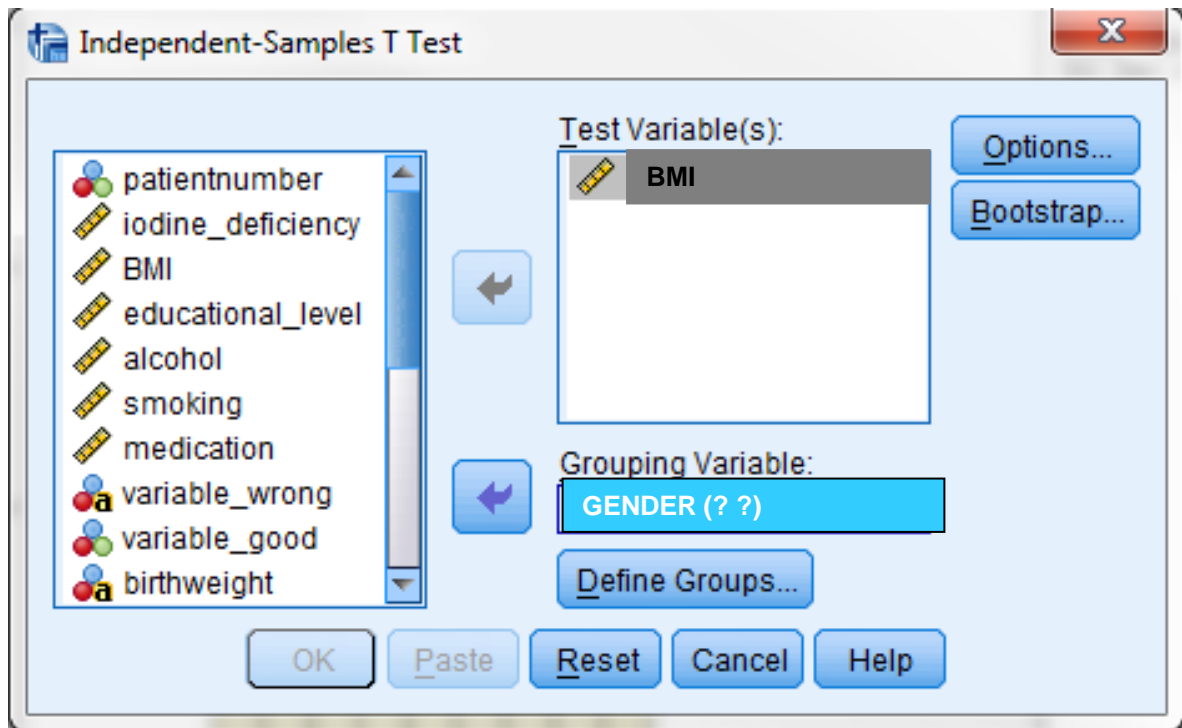
- Onderzoeksvraag:
Is er een significant verschil in gemiddelde Body Mass Index tussen mannen en vrouwen?
- Twee variabelen:
 - onafhankelijke variabele = man / vrouw
 - afhankelijke variabele = Body Mass Index

Independent-samples t-test - Graphical



T-Test – SPSS

- Analyze – Compare means – Independent samples t-test



T-Test – SPSS

- Aanname: Gelijke variantie tussen groepen

Levene's Test

$p < 0.05 \rightarrow$ twee groepen hebben geen vergelijkbare variantie (en standaard deviatie)

$p > 0.05 \rightarrow$ groep varianties zijn vergelijkbaar

Group Statistics

Status	N	Mean	Std. Deviation	Std. Error Mean
LN_BMI women	118	3,2409	,19568	,01801
LN_BMI men	98	3,1676	,15822	,01598

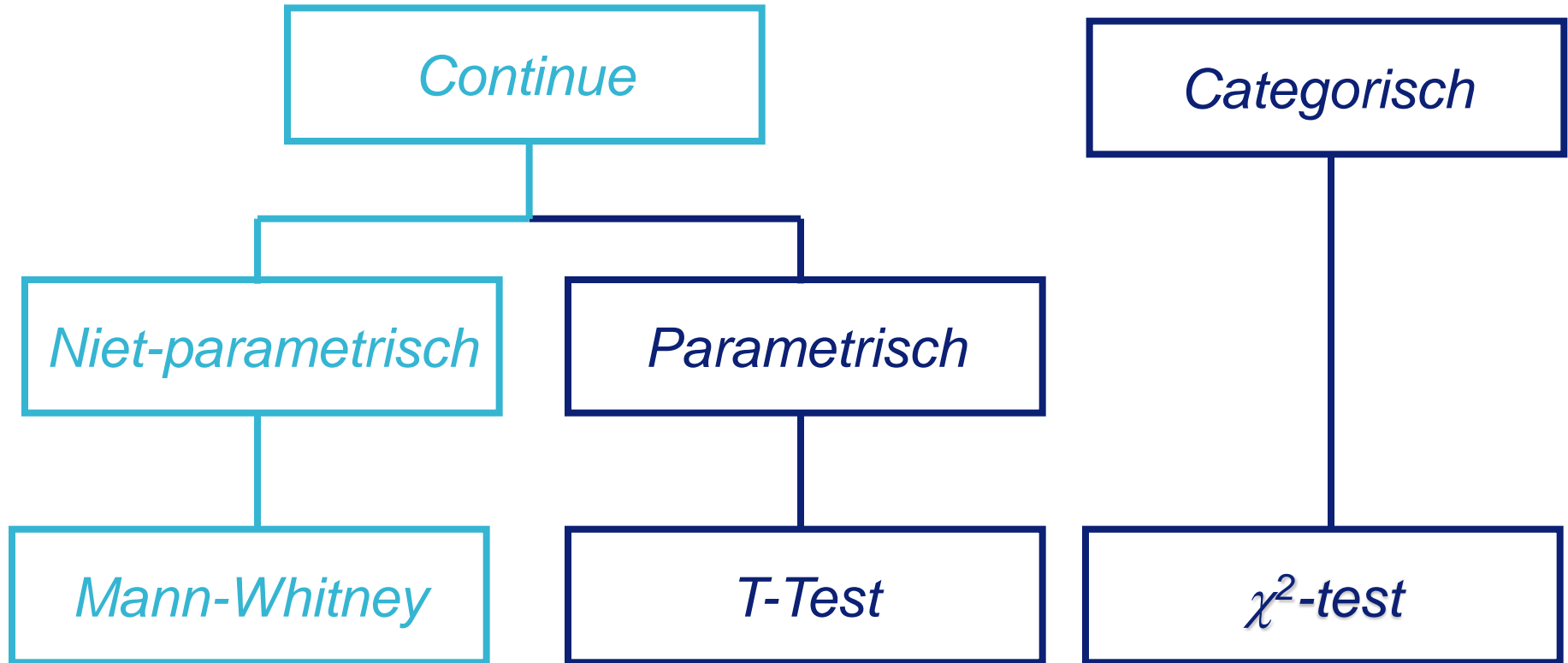
Independent Samples Test

	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	
								Lower	Upper
LN_BMI Equal variances assumed	2,285	,132	2,003	211	,003	,07326	,02456	,02486	,12166
LN_BMI Equal variances not assumed			3,042	213,858	,003	,07326	,02408	,02579	,12073

T-Test – SPSS – Conclusie

- Op basis van de test kunnen we concluderen:
- Ja, er is een verschil tussen de Body Mass Index tussen mannen en vrouwen in deze studie ($p = 0,003$)
- Vrouwen hebben gemiddeld een hoger BMI dan mannen.

Vergelijking twee medianen



Wanneer niet-parametrisch statistiek?

- Het is bekend dat de variabele een niet normale populatie distributie heeft
- De vorm of van de populatie distributie van de variabele is onbekend, en 'Normality tests' op de steekproef zijn significant
- Zeer kleine steekproeven
- Distributievrij (mean ranks)

Mann Whitney U Test

- Niet parametrische alternatief voor t-test
- Onderzoeksvraag: Is er een significant verschil in de variabele tussen de groepen?
- Twee variabelen:
 - één categorische, onafhankelijke variable (groep)
 - één continue, afhankelijke variabele (ranked)
- Test vertelt of er een statistisch significant verschil in de gemiddelde ranking is voor de twee groepen

Ranking

41	44	49	50	51
52	58	61	63	66
66	66	67	67	68
70	71	73	74	79
80	84	88	91	98

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Ranking

41	44	49	50	51
52	58	61	63	66
66	66	67	67	68
70	71	73	74	79
80	84	88	91	98

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

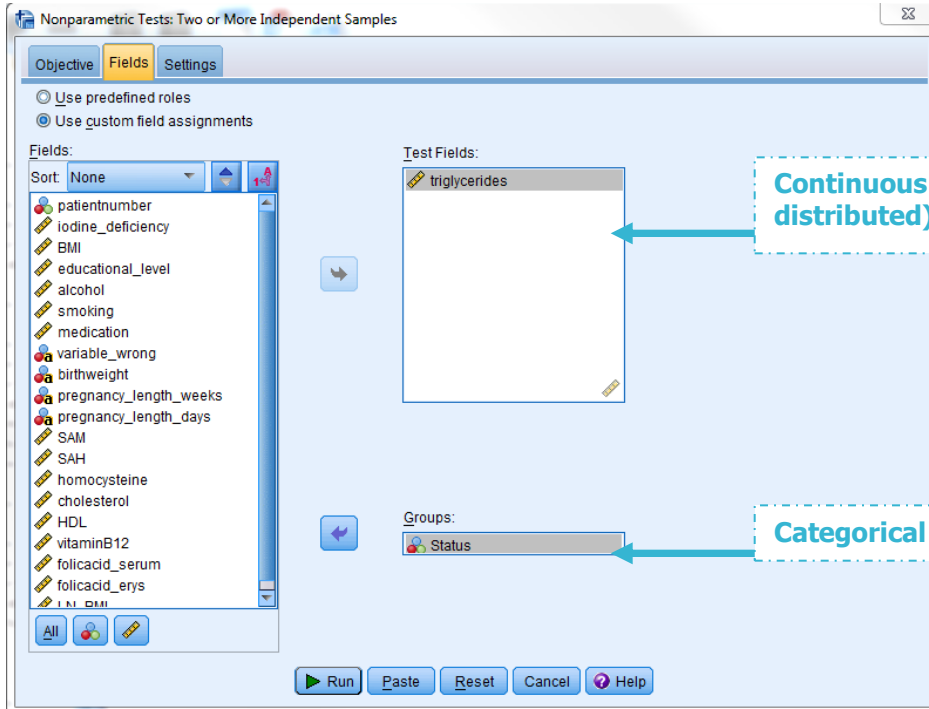
Mann Whitney U Test - voorbeeld

- Onderzoeksvraag: Is er een significant verschil in triglyceride levels tussen cases en controles?

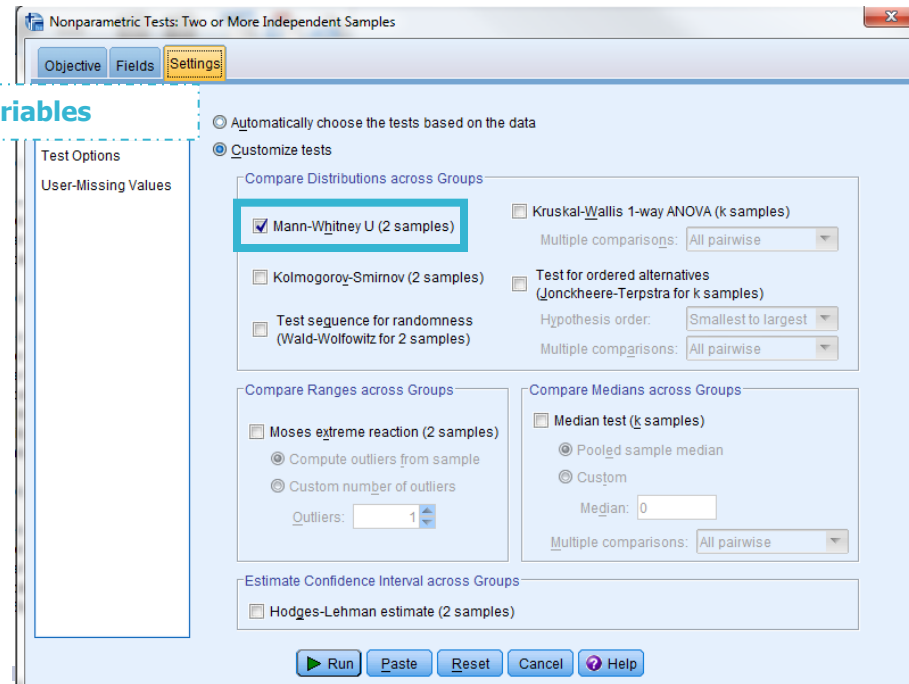
- Twee variabelen:
 - onafhankelijke variabele = case / control
 - afhankelijke variabele = triglyceride

Mann-Whitney U – SPSS

- Analyze – Non-Parametric tests – Independent samples



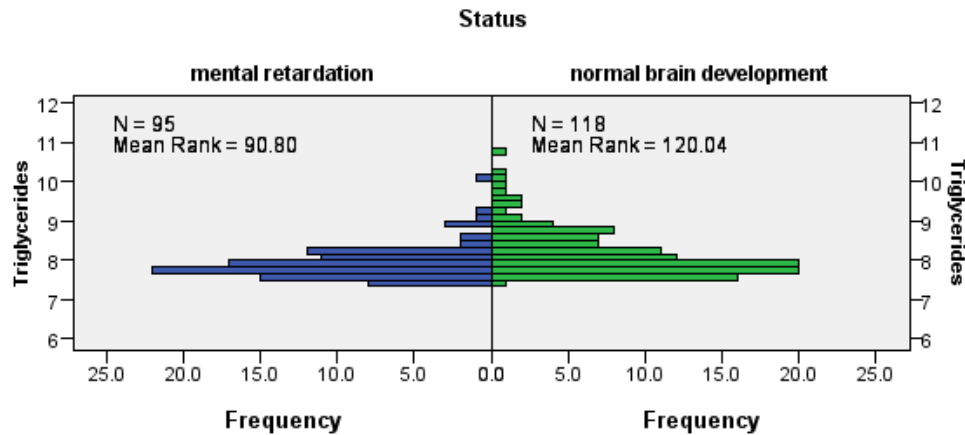
The dialog box shows the 'Fields' tab. Under 'Use custom field assignments', the 'Test Fields' list contains 'triglycerides'. The 'Groups' list contains 'Status'. A blue dashed box with an arrow points to 'triglycerides' with the text 'Continuous (not normally distributed) variables'. Another blue dashed box with an arrow points to 'Status' with the text 'Categorical variables'.



The dialog box shows the 'Settings' tab. Under 'Automatically choose the tests based on the data', the 'Customize tests' radio button is selected. In the 'Compare Distributions across Groups' section, the 'Mann-Whitney U (2 samples)' checkbox is checked and highlighted with a blue box. Other options include 'Kruskal-Wallis 1-way ANOVA (k samples)', 'Kolmogorov-Smirnov (2 samples)', and 'Test sequence for randomness (Wald-Wolfowitz for 2 samples)'. The 'Compare Ranges across Groups' section has 'Moses extreme reaction (2 samples)' selected. The 'Compare Medians across Groups' section has 'Median test (k samples)' selected. The 'Estimate Confidence Interval across Groups' section has 'Hodges-Lehman estimate (2 samples)' selected.

Mann-Whitney U – SPSS

Independent-Samples Mann-Whitney U Test

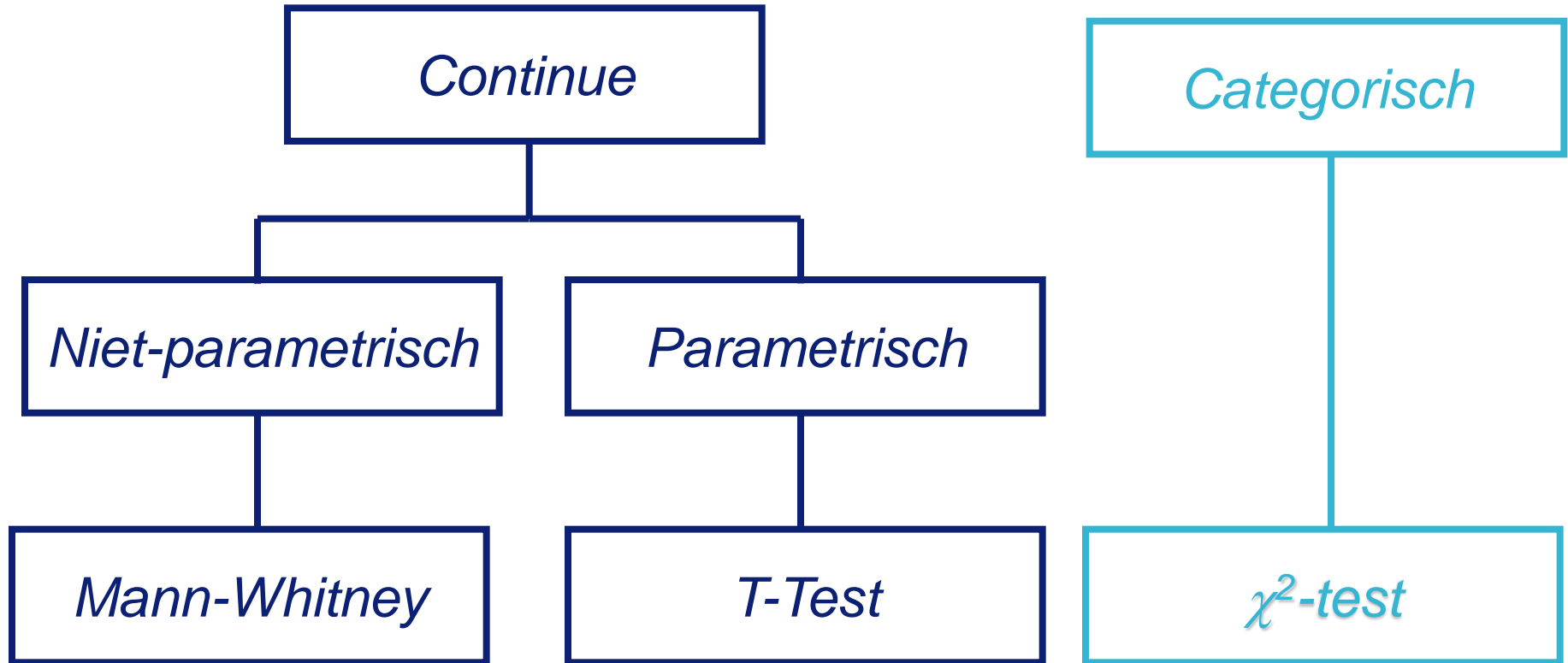


Total N	213
Mann-Whitney U	4,066.000
Wilcoxon W	8,626.000
Test Statistic	4,066.000
Standard Error	447.072
Standardized Test Statistic	-3.442
Asymptotic Sig. (2-sided test)	.001

P-value

Vergelijken van twee percentages

Vergelijking twee percentages



Chi-square test

- Onderzoeksvraag: Is er een significant verschil in proporties tussen de groepen?
- Twee variabelen:
 - een categorische, onafhankelijke variabele (groep)
 - een categorische, afhankelijke variabele
- $H_0: \pi_1 = \pi_0$
 $H_1: \pi_1 \neq \pi_0$
- Test vertelt of er een statistisch significant verschil is in proporties of fracties, kansen, percentages tussen de twee groepen

Chi-square test - Example

- Onderzoeksvraag: Is er een significant verschil in mortaliteit tussen patiënten gerandomiseerd naar actieve behandeling of placebo?
- Twee variabelen:
 - onafhankelijke variabele = therapie (0/1)
 - afhankelijke variabele = dood(0/1)
- $H_0: \pi_1 = \pi_0$
 $H_1: \pi_1 \neq \pi_0$

Voorbeeld- Hypothese

- H0 Mortaliteit in patiënten gerandomiseerd naar actieve behandeling = mortaliteit in patiënten gerandomiseerd naar placebo

- H1 Mortaliteit in patiënten gerandomiseerd naar actieve behandeling \neq mortaliteit in patiënten gerandomiseerd naar placebo

Voorbeeld– 2*2

	Death	Alive	
Active	50	450	500
Placebo	200	800	1000
	250	1250	1500

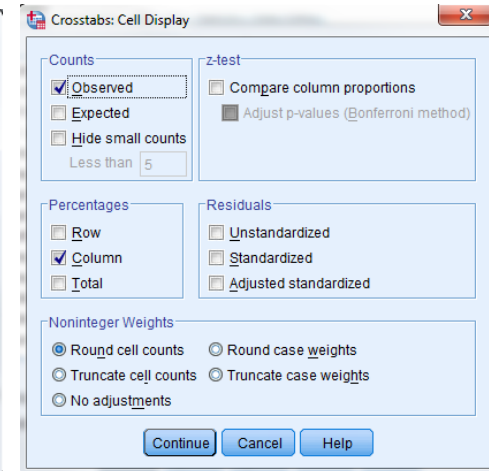
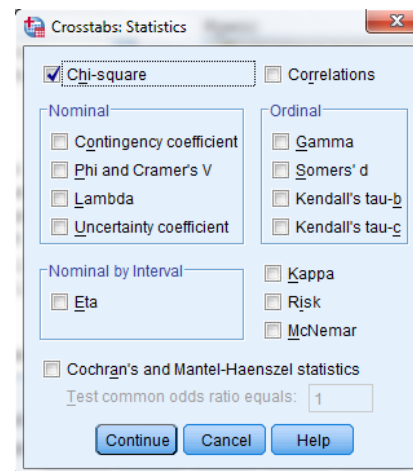
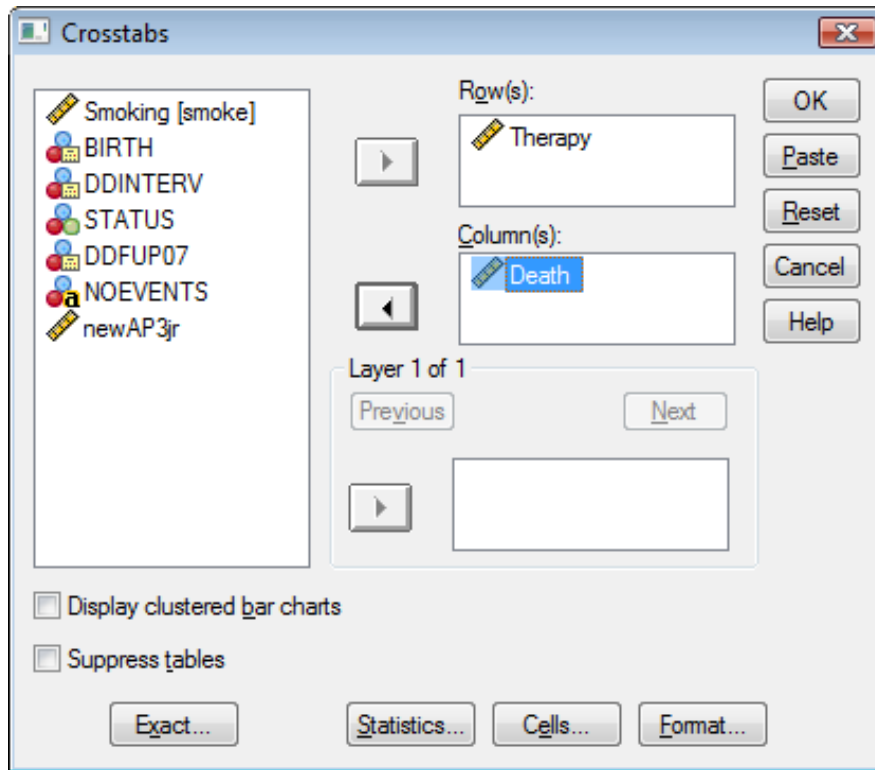
$$p_1 = X_1 / n_1 = 50 / 500 = 0.10$$

$$p_0 = X_0 / n_0 = 200 / 1000 = 0.20$$

$$p = X / n = 250 / 1500 = 0.17$$

SPSS – Chi-square test

- Analyze – Descriptive Statistics – Crosstabs



SPSS – Chi-square test

Crosstabs

[DataSet1] E:\t.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Therapy * Death	1500	100,0%	0	,0%	1500	100,0%

Therapy * Death Crosstabulation

			Death		Total
			,00 Alive	1,00 Death	
Therapy	,00 Placebo	Count	800	200	1000
		% within Therapy	80,0%	20,0%	100,0%
		% within Death	64,0%	80,0%	66,7%
1,00 Active	Count	450	50	500	
	% within Therapy	90,0%	10,0%	100,0%	
	% within Death	36,0%	20,0%	33,3%	
Total	Count	1250	250	1500	
	% within Therapy	83,3%	16,7%	100,0%	
	% within Death	100,0%	100,0%	100,0%	

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	24,000 ^a	1	,000		
Continuity Correction ^b	23,285	1	,000		
Likelihood Ratio	25,796	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	23,984	1	,000		
N of Valid Cases	1500				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 83,33.

SPSS – Chi-square test - Conclusie

- Op basis van de test kunnen we concluderen:
- Ja, er is een significant verschil in mortaliteit tussen patiënten behandeld met placebo en actieve therapy ($p < 0.001$)

Samenvatting vergelijken groepen

