Onderzoeksmethoden: Statistiek 3

Peter de Waal
(gebaseerd op slides Peter de Waal, Marjan van den Akker)

Departement Informatica
Beta-faculteit, Universiteit Utrecht

Vandaag

- Recap
- Centrale limietstelling
- T-verdeling
- Toetsen van hypotheses

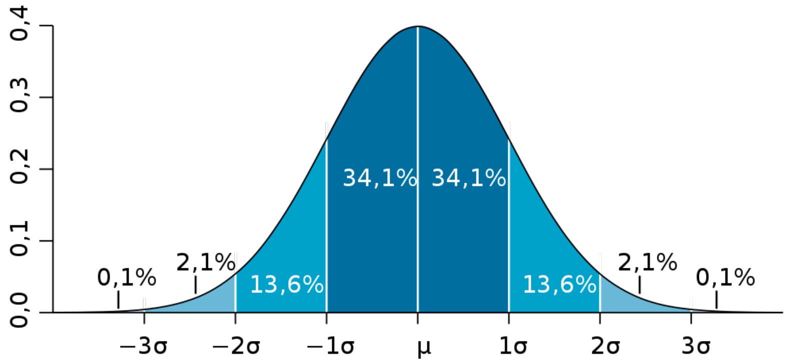
Recap 1

Begrippen

- Experiment
- Stochastische variabele
- Theoretische kansverdelingen
- Verwachting ("gemiddelde"), variantie, standaarddeviatie...
- Z-scores

Recap 2

- Normale verdeling: $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
- Standaardnormale verdeling: speciaal geval met $\mu = 0$, $\sigma = 1$.



Reality check

Q: Wat als scores uit de “echte wereld” helemaal niet lijken op een theoretische verdeling?

Q: Wat als we de populatieparameters niet kennen (b.v. μ en σ?) Trek een aselecte steekproef $X_1, X_2, \ldots, X_n$ van omvang n en gebruik:

Steekproefgemiddelde

$$\overline{X}(n) = \frac{\sum_{i=1}^{n} X_i}{n}$$

Het steekproefgemiddelde, zoals in bovenstaande formule is ook een *stochastische variabele* en heeft dus een verdeling, verwachting, variantie, …

Centrale limietstelling

- Steekproef $X_1, \ldots, X_n$: elke afzonderlijke X_i is onafhankelijk en identiek verdeeld, gemiddelde μ en variantie σ^2.
- $\overline{X}(n) = (X_1 + X_2 + \cdots + X_n)/n$

Stelling

Als $n \to \infty$, dan is $\overline{X}(n)$ bij benadering normaal verdeeld met gemiddelde μ en variantie σ^2/n.

- $\dfrac{\sigma}{\sqrt{n}}$ heet de **standaardfout** (**van het gemiddelde**)
- $Z(n) = \dfrac{\overline{X}(n) - \mu}{\left(\sigma/\sqrt{n}\right)}$ is standaardnormaal verdeeld.
- In de praktijk is $n \geq 30$ voldoende.

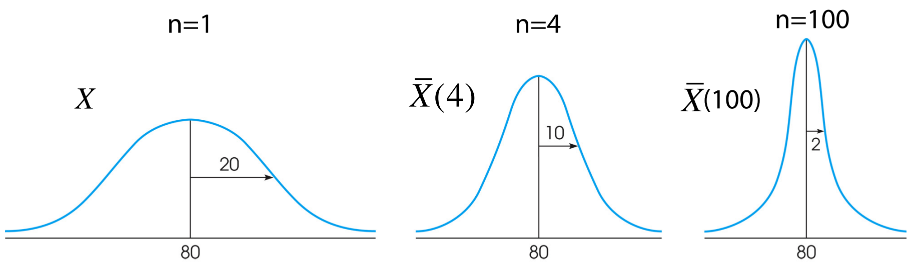
Centrale limietstelling: wat betekent dit?

- **Steekproefverdeling**:
  - Meet individuen op variabele X.
  - Herhaal dit m keer (b.v. $m = 1000$) en noteer individuele waarden.
  - Histogram representeert **steekproefverdeling** (Eng: **sample distribution**)
- **Verdelingen van steekproefgemiddelde**
  - Neem een steekproef van n individuen en meet en bepaal steekproefgemiddelde $\overline{X}(n)$.
  - Herhaal dit m keer (b.v. $m = 1000$) en noteer de afzonderlijke steekproefgemiddelden.
  - Histogram representeert de **verdeling van het steekproefgemiddelde** (Eng: **sampling distribution**)

Voorbeeld 1

Voorbeeld

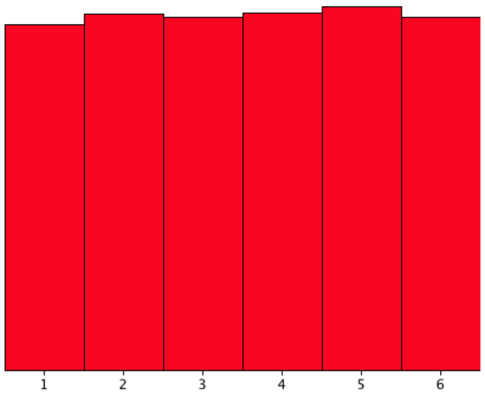
- Gemeten variabele: gewicht X
- Normale verdeling ($\mu = 80$ kg, $\sigma = 20$ kg).



Voorbeeld 2

Voorbeeld: worp van één dobbelsteen:

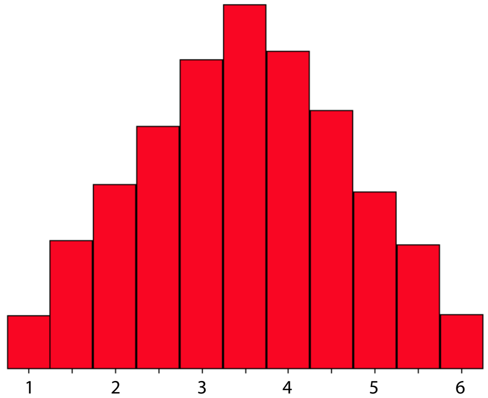
- Elke steekproef = 1 worp:
  histogram: tel aantal ogen (*sample distribution*)



Voorbeeld 2

Voorbeeld: worp van twee dobbelstenen:

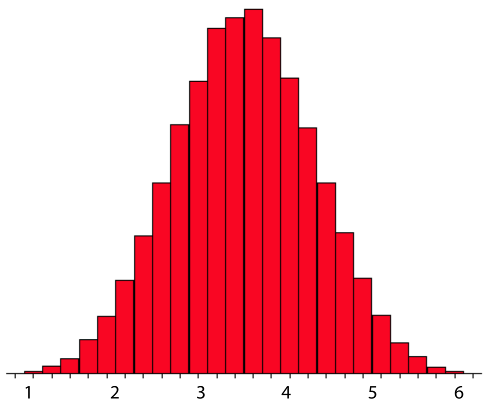
- Elke steekproef = 2 worpen:
  histogram: tel gemiddeld aantal ogen
  (is dus *sampling distribution bij steekproefgrootte* $n = 2$)



Voorbeeld 2

Voorbeeld: worp van vijf dobbelstenen:

- Elke steekproef = 5 worpen:
  histogram: tel gemiddeld aantal ogen
  (*sampling distribution voor* $n = 5$)



Back to reality

$$Z(n) = \frac{\overline{X}(n) - \mu}{\left(\sqrt{\sigma^2/n}\right)} = \frac{\overline{X}(n) - \mu}{(\sigma/\sqrt{n})}$$

Is dus bij benadering *standaard normaal verdeeld* (mits $n \geq 30$).

Leuk! maar meestal weten we σ helemaal niet...

Oplossing: vervang σ door schatting, n.l. steekproefvariantie s.

$$t(n-1) = \frac{\overline{X}(n) - \mu}{\left(\sqrt{s^2/n}\right)} = \frac{\overline{X}(n) - \mu}{(s/\sqrt{n})}$$

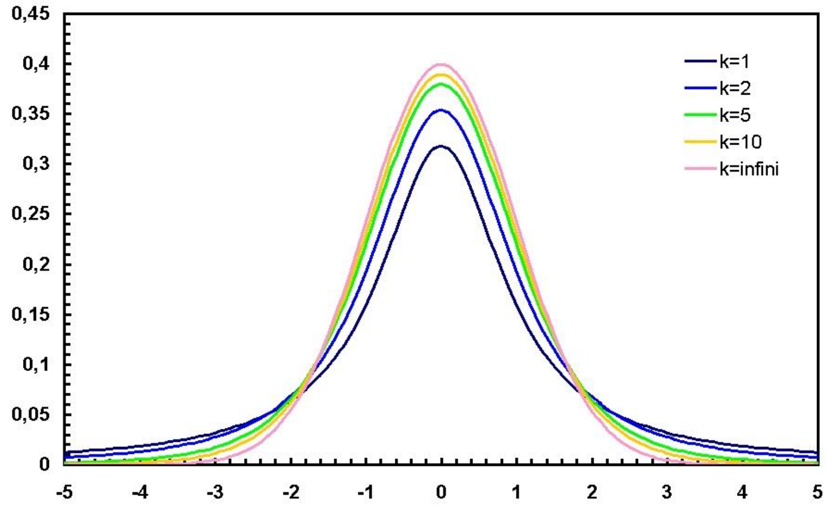
$t(n-1)$ heeft *geen* standaardnormale verdeling, maar een zgn. t-verdeling met $n-1$ vrijheidsgraden.

t-verdeling

- Is een *familie* van verdelingen
- Lijkt op de standaardnormale verdeling in vorm en spreiding
- Heeft iets “meer massa” in de staarten (platter)
- Heeft één parameter: *aantal vrijheidsgraden* (*Eng: degrees of freedom (df)*)
- Voor $df \to \infty$ (praktisch: $df \geq 120$) is de t-verdeling gelijk aan de standaardnormale verdeling
- Wordt soms ook *Student verdeling* genoemd.

William Sealey Gosset (1876–1937)

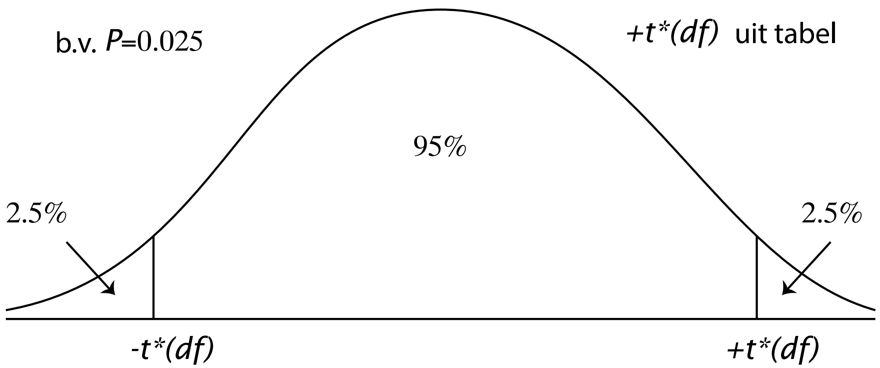
t-verdeling: plots



(Source: wikipedia)

t-verdeling: kansen

- $t(n) = \dfrac{\overline{X}(n) - \mu}{(s/\sqrt{n})}$ heeft t-verdeling met $n-1$ vrijheidsgraden.
- Een aantal kansen kunnen in tabellen opgezocht worden:



t-verdeling: speciale waarden

Tabellen voor speciale kansen:

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	**0.50**	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
two-tails	**1.00**	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850

t-verdeling: speciale waarden

Tabellen voor speciale kansen:

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$
one-tail	**0.50**	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**
two-tails	**1.00**	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**
df						
1	0.000	1.000	1.376	1.963	3.078	6.314
2	0.000	0.816	1.061	1.386	1.886	2.920
3	0.000	0.765	0.978	1.250	1.638	2.353
4	0.000	0.741	0.941	1.190	1.533	2.132
5	0.000	0.727	0.920	1.156	1.476	2.015
6	0.000	0.718	0.906	1.134	1.440	1.943
7	0.000	0.711	0.896	1.119	1.415	1.895
8	0.000	0.706	0.889	1.108	1.397	1.860
9	0.000	0.703	0.883	1.100	1.383	1.833
10	0.000	0.700	0.879	1.093	1.372	1.812
11	0.000	0.697	0.876	1.088	1.363	1.796

Toetsen

- Variabele X: aantal uren gamen per week
- Steekproef van $n = 10$ UU informatica studenten:
  - 18, 25, 28, 21, 23, 18, 18, 26, 25, 21
- We willen nu een uitspraak doen over het gemiddelde over de hele populatie van alle UU informatica studenten
- **Procedure**:
  1. Formuleer hypotheses
  2. Kies teststatistiek en leg criterium vast
  3. Bereken teststatistiek uit steekproef
  4. Neem beslissing over hypothese

Toetsen

1. Formuleer hypotheses

- Gemiddelde over andere universiteiten is bekend $\mu = 20$
- We vermoeden dat dit in Utrecht afwijkt

Twee hypotheses:

- Nulhypothese $H_0 : \ \mu = 20$
- Alternatieve hypothese $H_1 : \ \mu \neq 20$

Let op:

- Research-hypothese = alternatieve hypothese (H_1)!
- Hypotheses worden geformuleerd voor populatieparameters.
- Nulhypothese en alternatieve hypothese zijn:
  - Uitsluitend
  - Uitputtend

Toetsen (t-toets voor één gemiddelde)

2. Kies teststatistiek en leg criterium vast

- Test statistiek: steekproefgemiddelde $\overline{X}$
- Als H_0 waar is, wat zijn dan aannemelijke waarden voor $\overline{X}$?
  - $\overline{X} = 20$?
  - $\overline{X} = 18$?
  - $\overline{X} = 26$?
- Wat zijn dan aannemelijke waarden voor $t = \dfrac{\overline{X} - 20}{(s/\sqrt{n})}$?
  - $t = 0$?
  - $t = -1$?
  - $t = +5$?

Toetsen (t-toets voor één gemiddelde)

2a. Kies teststatistiek

- $H_0: \ \mu = 20$
  $H_1: \ \mu \neq 20$
- Neem aan:
  - Waarden in de steekproef zijn onafhankelijk, identiek verdeeld en normaal verdeeld.
- Als H_0 waar is, dan geldt dat

$$t = \frac{\overline{X}(10) - 20}{(s/\sqrt{10})} \quad \text{met} \quad s^2 = \frac{SS}{df} = \frac{\sum_{i=1}^{10}(X_i - \overline{X})^2}{9}$$

heeft t-verdeling met $df = 9$ vrijheidsgraden.

Toetsen (t-toets voor één gemiddelde)

2B. Leg criterium vast

- Kies tevoren een gewenst *significantieniveau* α. (bv $\alpha = 0.05$)
- Bepaal *verwerpingsgebied* (of *kritieke gebied*):

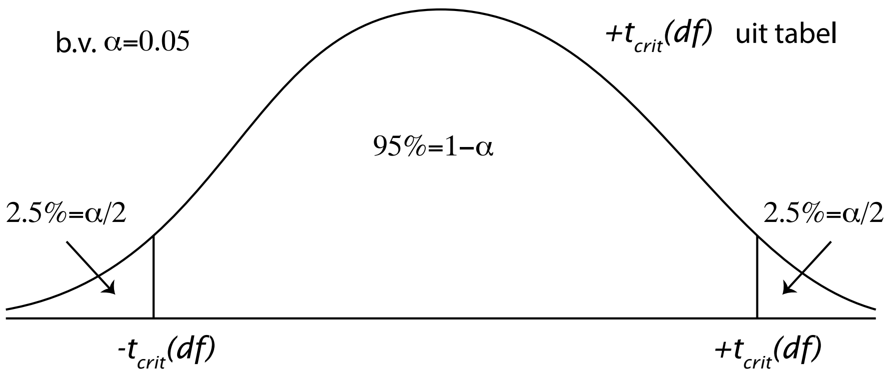
Beslissingregel:

- Verwerp H_0, als $t \leq -t_{crit}$ of $t \geq +t_{crit}$
- Verwerp H_0 niet, als $-t_{crit} < t < +t_{crit}$

Hoe kiezen we t_{crit}?

- We willen bijvoorbeeld $\alpha = 0.05$, ofwel de kans dat we H_0 verwerpen terwijl deze waar is, is 5%.
- Kritieke gebied = gebied waarvoor kans op een waarneming onder H_0 gelijk is aan α.

Toetsen (t-toets voor één gemiddelde)



Toetsen (t-toets voor één gemiddelde)

Leg criterium vast

- We willen $\alpha = 0.05$, ofwel de kans dat we H_0 verwerpen terwijl hij waar is, is 5%.

Criterium wordt: Kies t_{crit} zodat:

- $P(t \geq t_{crit}) = \alpha/2 = 0.025$ en $P(t \leq -t_{crit}) = \alpha/2 = 0.025$ voor t-verdeling met $df = 9$ vrijheidsgraden,
- dus $\alpha/2 = 0.025$ "aan iedere kant" ("two-tails" $\alpha = 0.05$)
- met tabel: $t_{crit}(df = 9) = ??$

Tabel voor de t-verdeling

***t* Table**

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850

Toetsen (t-toets voor één gemiddelde)

3. Bereken teststatistiek uit steekproef

- Bereken uit steekproef:
  Het steekproefgemiddelde $\overline{X} = 22.3$ en $s = 3.65$, $\mu = 20$, en $n = 10$.
  Bepaal de waarde van toetsinggrootheid.

$$t = \frac{\overline{X}(n) - \mu}{(s/\sqrt{n})} = \frac{22.3 - 20.0}{(3.65/\sqrt{10})} = \frac{2.3}{1.155} = 1.991$$

- $t_{crit}(df = 9) = 2.262$ dus

4. Neem beslissing? H_0 niet verwerpen

Enkelzijdig/tweezijdig toetsen

- Bij welk steekproefgemiddelde zijn we het eens met H_0?
- Tweezijdig toetsen:
  - Hypothese H_0 : $\mu = 20$
  - Alternatieve hypothese H_1 : $\mu \neq 20$
- Eenzijdig toetsen:
  - Hypothese H_0 : $\mu \leq 20$
  - Alternatieve hypothese H_1 : $\mu > 20$
- Deze keuze moet je maken vóór je data gaat verzamelen!

Eenzijdige t-toets voor één gemiddelde

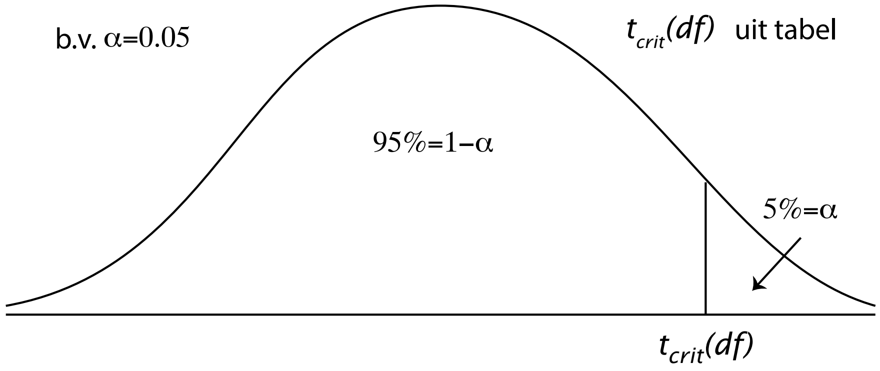
1. Formuleer hypotheses

- Eenzijdig toetsen:
  - H_0 : $\mu \leq 20$ (aantal uren gamen per week)
    H_1 : $\mu > 20$
  - Nog steeds geldt:

$$t = \frac{\overline{X}(n) - \mu}{(s/\sqrt{n})}$$

heeft t-verdeling met $df = 9$ vrijheidsgraden.

Student's t-verdeling



Eenzijdige t-toets voor één gemiddelde

2. Kies teststatistiek en leg criterium vast

Nu beslissingsregel van de vorm:

- Verwerp H_0, indien $t_{obs} \geq t_{crit}$
- Verwerp H_0 niet, indien $t_{obs} < t_{crit}$
- We willen de kans dat we H_0 verwerpen terwijl hij waar is is 5%. (Significantieniveau α)
- Criterium wordt:
  Kies t_{crit} zodat $P(t \geq t_{crit}) = \alpha = .05$ voor t-verdeling met $df = 9$ vrijheidsgraden Tabel: $t_{crit}(df = 9) = ???$

Tabel voor de t-verdeling

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	**0.50**	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
two-tails	**1.00**	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	[illegible]	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	$\boxed{1.833}$	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	[illegible]	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850

Toetsen (t-toets voor één gemiddelde, enkelzijdig)

3. Bereken teststatistiek uit steekproef

- Bereken uit steekproef:
  Het steekproefgemiddelde $\overline{X} = 22.3$ en $s = 3.65$, $\mu = 20$, en $n = 10$.
  Bepaal de waarde van toetsinggrootheid.

$$t = \frac{\overline{X}(n) - \mu}{(s/\sqrt{n})} = \frac{22.3 - 20.0}{(3.65/\sqrt{10})} = \frac{2.3}{1.155} = 1.991$$

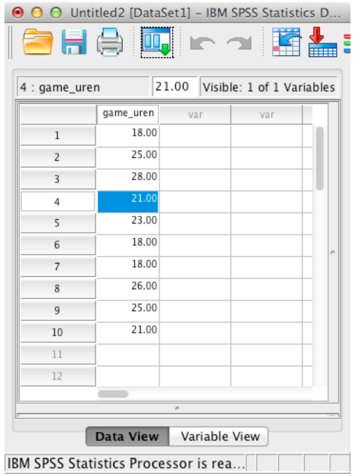
- $t_{crit}(df = 9) = 1.833$. Nu is $t_{obs} > t_{crit}$, dus

4. Neem beslissing? H_0 wel verwerpen

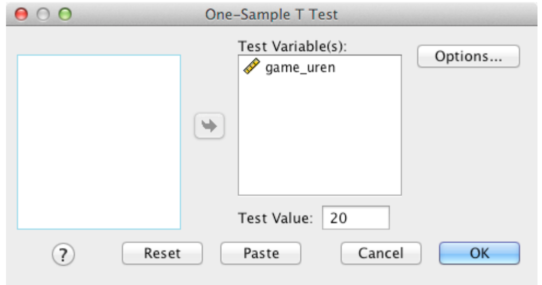
Samenvatting

	H_0	H_1	H_0 verwerpen als
Tweezijdig	$\mu = \mu_0$	$\mu \neq \mu_0$	$t \leq -t_{crit}$ of $t \geq t_{crit}$
Eenzijdig (rechts)	$\mu \leq \mu_0$	$\mu > \mu_0$	$t \geq t_{crit}$
Eenzijdig (links)	$\mu \geq \mu_0$	$\mu < \mu_0$	$t \leq -t_{crit}$

t-toets in SPSS



Menu: Analyze $\rightarrow$ Compare Means $\rightarrow$ One Sample T Test



t-toets in SPSS: Output

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
game_uren	10	22.3000	3.65300	1.15518

One-Sample Test

	Test Value = 20					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
game_uren	1.991	9	.078	2.30000	-.3132	4.9132

p-waarde of significantie

De p-waarde of significantie ("Sig." in SPSS) van een gegeven steekproef uitkomst is

- de kans dat onder de nulhypothese de geobserveerde waarde van de toetsingsgrootheid wordt behaald of overschreden.
- Geeft aan hoe extreem de waarde van de toetsingsgrootheid is.

Dus:

- t: toetsingsgrootheid
- t_{obs}: waarde van de toetsingsgrootheid uit steekproef.
- p-waarde= $P(|t(df)| \geq |t_{obs}|)$, dus tweezijdig!
- Eenzijdig: waarschijnlijkheid van geobserveerde waarde is *helft* van de p-waarde!

p-waarde of significantie (2)

- Hoe kleiner de p-waarde hoe extremer de uitkomst.
- p-waarde kleiner dan gegeven grens (bijv 5%), dan significante uitkomst ofwel significant verschil met H_0, dus H_0 wordt verworpen.

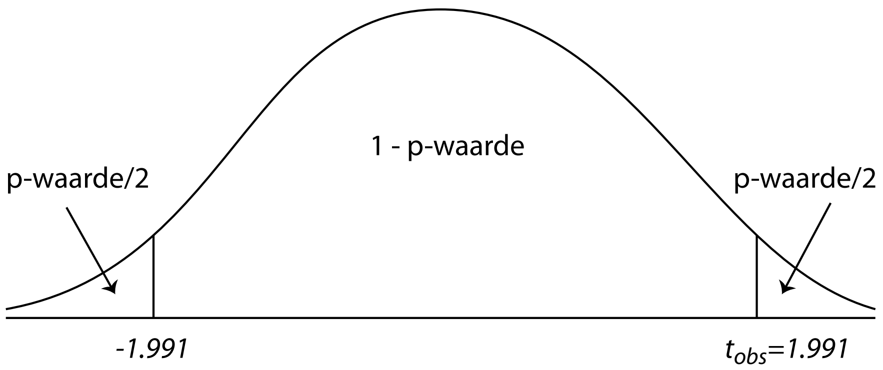
One-Sample Test

	Test Value = 20					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
game_uren	1.991	9	.078	2.30000	-.3132	4.9132

p-waarde of significantie: voorbeeld

- *Tweezijdige* t-toets aantal uren gamen per week UU informatica studenten
  - Hypothese H_0: $\mu = 20$
  - Alternatieve hypothese H_1: $\mu \neq 20$
- Voorbeeld steekproef:
  - 18, 25, 28, 21, 23, 18, 18, 26, 25, 21:
  - $t_{obs} = \dfrac{2.3}{1.155} = 1.991$
- Mbv. Excel T.DIST.2T(1.991, df=9)= 0.0776
- $0.077 > 0.05$ dus geen significante afwijking van H_0

Student's t-verdeling tweezijdig



Fouten

What could possily go wrong?

		Toestand	
		H_0 waar	H_1 waar
Beslissing	niet verwerpen H_0	OK Kans $= 1 - \alpha =$ betrouwbaarheid	Type II fout
	verwerp H_0	Type I fout Kans $= \alpha =$ significantieniveau	OK

Check!

Q: Als je het significantieniveau α verkleint, dan neemt het kritieke gebied toe?

A: Fout.

Q: Als je het significantieniveau α verkleint, dan neemt de kans op een type I fout toe?

A: Fout.

t-toets voor één gemiddelde

- ENG: One-sample t-test.
- Vergelijking van één steekproefgemiddelde met een 'norm' (een van te voren bepaald gemiddelde).
- σ uit populatie is niet bekend en het steekproefaantal is klein ($n < 120$).
- Meetwaarden onafhankelijk en identiek normaal verdeeld (met zelfde gemiddelde en variantie).
- Oplossing: t-verdelingen (vrijheidsgraden spelen een rol)

$$t_{obs} = \frac{\overline{X}(n) - \mu}{(s/\sqrt{n})}$$

- t-table: zie boek, online, Excel, calculator...

Betrouwbaarheidsinterval

- Stel: ik meet steekproefgemiddelde $\overline{X}(n) = 23.4$
- Kan ik nu met 95% betrouwbaarheid zeggen in welk gebied het onbekende populatiegemiddelde μ ligt?
- 95% betrouwbaarheidsinterval:

$$\left[\overline{X}(n) - \frac{s}{\sqrt{n}} \times t_{crit} \quad ; \quad \overline{X}(n) + \frac{s}{\sqrt{n}} \times t_{crit}\right]$$

waarbij t_{crit} de kritieke waarde is voor $\alpha = 0.05$ bij $df = n - 1$.

Betrouwbaarheidsinterval: voorbeeld

Student	Rick	Mark	Tom	Ken	Edwin
# uren gamen per week	18	24	24	21	18

- $\overline{X} = 21$
- $n = 5$
- $s^2 = 9$
- $t_{crit}(df = 4) = 2.776$ voor $\alpha = 0.05$
- 95% betrouwbaarheidsinterval

$$\left[21 - \frac{3}{\sqrt{5}} \times 2.776 \; ; \; 21 + \frac{3}{\sqrt{5}} \times 2.776\right] = \left[17.28 \; ; \; 24.72\right]$$

Betrouwbaarheidsinterval (Interpretatie)

- $(1-\alpha)100\%$ betrouwbaarheidsinterval, b.v. $\alpha = 0.05$.
- NIET:
  De echte waarde van μ valt met kans 95% binnen het interval
- WEL:
  Als we vaker een steekproef nemen met 5 studenten en telkens de bijbehorende betrouwbaarheidsintervallen uitrekenen, dan valt de echte waarde van μ binnen 95% van deze intervallen.

Betrouwbaarheidsinterval $\leftrightarrow$ Toetsen

- $H_0 : \mu = \mu_0$ wordt verworpen in tweezijdige t-toets met significantie α

dan en slechts dan als

- μ_0 ligt buiten het $(1-\alpha) * 100\%$ betrouwbaarheidsinterval

Voorbeeld in SPSS:

One-Sample Test

	Test Value = 20					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
game_uren	1.991	9	.078	2.30000	-.3132	4.9132

Speciaal geval t-toets: toetsen van correlatie

Onderzoeksvermoeden: Er bestaat een correlatie tussen aantal geconsumeerde blikjes cola en regels gemaakte code:

Variabelen:

- X: aantal geconsumeerde blikjes cola
- Y: aantal geproduceerde regels code

Dit kan getoetst worden met een t-toets:

1. Formuleer hypothese

- $H_0 : \rho = 0$ (ρ = correlatie-coëfficient cola-code *in populatie*)
- $H_1 : \rho \neq 0$.

Toetsen van correlatie (2)

2. Kies teststatistiek en leg criterium vast

- Kies $\alpha = 0.05$.
- Gebruik de volgende t-statistiek:

$$t = r \times \sqrt{\frac{n-2}{1-r^2}}$$

  met
- r = correlatie cola-code *in steekproef*
- n = steekproefgrootte
  Eigenschap:
- t heeft t-verdeling met $n-2$ vrijheidsgraden.

Toetsen van correlatie (3)

3. Bereken teststatistiek uit steekproef

In College 1 hadden we een steekproef met:

- $n = 5$
- $r = 0.9995$
  Dus

$$t = 0.9995 \times \sqrt{\frac{3}{1 - 0.9995^2}} = 54.75$$

$$t_{crit}(df = 3) = 3.182$$

Dus:

4. Neem beslissing: Verwerp H_0. De correlatie tussen cola en code wijkt significant af van 0.

Gedaan:

- Centrale Limietstelling
- T-verdeling
- Toetsen van hypothese: One-sample T-toets.

Volgende week:

- Meer T-toetsen
- Chi-kwadraat toetsen