Hypothesetoetsen voor twee populaties

Sandra Van Aert

24 november 2011

Inleiding

- doel: gelijkenissen/verschillen tussen 2 populaties ontdekken en meten
- meest algemeen: cumulatieve verdelingsfuncties vergelijken

$$H_0: F_1 = F_2 \text{ versus } H_a: F_1 \neq F_2$$

 meestal beperkt men zich tot het vergelijken van gemiddeldes, varianties, medianen, proporties,...

Voorbeelden

populatiegemiddeldes

$$H_0: \mu_1 = \mu_2$$
 versus $H_a: \mu_1 > \mu_2$
 $H_a: \mu_1 < \mu_2$
 $H_a: \mu_1 \neq \mu_2$

populatievarianties

$$H_0: \sigma_1^2 = \sigma_2^2 \text{ versus } H_a: \sigma_1^2 \leq \sigma_2^2$$

populatieproporties

$$H_0: \pi_1 = \pi_2 \text{ versus } H_a: \pi_1 \leq \pi_2$$

...

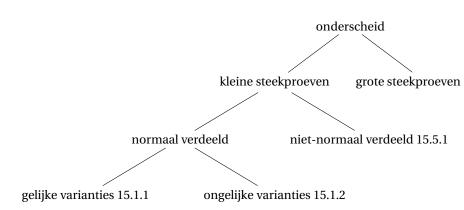
Belangrijk

gepaarde waarnemingen of afhankelijke steekproeven

1

niet-gepaarde waarnemingen of onafhankelijke steekproeven

Twee populatiegemiddeldesniet-gepaarde waarnemingen



Twee populatiegemiddeldesgepaarde waarnemingen

algemeen

$$H_0: \mu_1 = \mu_2$$
 versus $H_a: \mu_1 > \mu_2$ rechtseenzijdig $H_a: \mu_1 < \mu_2$ linkseenzijdig $H_a: \mu_1 \neq \mu_2$ tweezijdig

herschrijven als

$$H_0: \underbrace{\mu_1 - \mu_2}_{\delta} = 0 \text{ versus } H_a: \underbrace{\mu_1 - \mu_2}_{\delta} > 0 \text{ rechts}$$

$$H_a: \delta < 0 \text{ links}$$

$$H_a: \delta \neq 0 \text{ tweezijdig}$$

Steekproeven

steekproef 1		steekproef 2		verschilvariabele
X_{11}	_	X_{21}	=	D_1
X_{12}	_	X_{22}	=	D_2
X_{13}	_	X_{23}	=	D_3
:		:		:
X_{1n}	_	X_{2n}	=	D_n

Gelijkenis met één variabele

- we belanden dus in situatie met één variabele!
- toen:

$$\begin{cases} H_0: \mu = \mu_0 \\ H_a: \mu \underset{\neq}{\leq} \mu_0 \end{cases}$$

▶ nu:

$$\begin{cases} H_0: \delta = 0 \\ H_a: \delta \lesssim 0 \end{cases}$$

Toetsingsgrootheid

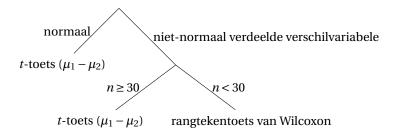
- vereist
 - \rightarrow normaal verdeelde verschilvariabele D
 - → of grote steekproeven

$$T = \frac{\overline{D} - 0}{s_D / \sqrt{n}} \sim t_{n-1} \quad \text{bij ongekende } \sigma_D$$

$$Arr Z = \frac{\overline{D} - 0}{\sigma_D / \sqrt{n}} \sim N(0, 1)$$
 bij gekende σ_D

- wat bij kleine steekproeven en niet-normaal verdeelde verschilvariabele?
 - → rangtekentoets van Wilcoxon (zie verder)

Beslissingsboom gepaarde waarnemingen



Voorbeeld

aantal bacteriën

```
        vóór
        na
        verschil

        1)
        6.98
        6.95
        0.03

        2)
        7.08
        6.94
        0.14

        ...
        ...
        ...
        ...

        12)
        7.69
        6.99
        0.70
```

- vóór en na dezelfde stalen⇒ gepaarde waarnemingen
- $\vec{d} = 0.2583$ $s_D = 0.3565$

toets of aantal bacteriën gedaald is

$$H_0: \mu_{\text{voor}} = \mu_{\text{na}}$$
 $H_a: \mu_{\text{voor}} > \mu_{\text{na}}$ $H_0: \mu_{\text{voor}} - \mu_{\text{na}} = 0$ $H_a: \mu_{\text{voor}} - \mu_{\text{na}} > 0$ $H_0: \delta = 0$ $H_a: \delta > 0$

• $t = \frac{0.2583}{0.3565/\sqrt{12}} = 2.5101$

• kritieke waarde: $t_{\alpha;11} = t_{0.05;11} = 1.7959$ $t > t_{\alpha;11} \Rightarrow \text{verwerp } H_0$

• $p\text{-waarde: } p = P(t_{11} > 2.5101) < 0.025$ R: "1-pt(t,11)" = 0.014491 Matlab: "1-tcdf(t,11)" $p < \alpha \Rightarrow \text{verwerp } H_0$

aantal bacteriën

	ii bacteriori	
	<u>vóór</u>	<u>na</u>
1)	6.98	6.95
2)	7.08	6.94
:	:	:
12)	7.69	6.99
	↓	↓
	$\overline{x}_1 = 6.7208$	$\overline{x}_2 = 6.4625$
	$s_1^2 = 0.7363$	$s_2^2 = 0.4341$

 stel dat u niet herkent dat het gaat om gepaarde waarnemingen, en doet alsof de steekproeven onafhankelijk zijn

toets of aantal bacteriën gedaald is

$$H_0: \mu_{\text{voor}} = \mu_{\text{na}}$$
 $H_a: \mu_{\text{voor}} > \mu_{\text{na}}$
 $H_0: \mu_{\text{voor}} - \mu_{\text{na}} = 0$ $H_a: \mu_{\text{voor}} - \mu_{\text{na}} > 0$

$$t = \frac{6.7208 - 6.4625}{\sqrt{\frac{0.7363}{12} + \frac{0.4341}{12}}} = 0.8272$$

- ► kritieke waarde: $t_{\alpha;\nu} = t_{0.05;21} = 1.7207$ $t < t_{\alpha;21} \Rightarrow \text{aanvaard } H_0$
- ► *p*-waarde: $p = P(t_v > 0.8272) = P(t_{21} > 0.8272) = 0.2088$ $p > α \Rightarrow \text{aanvaard } H_0$

Twee populatievarianties

- noodzakelijke tussenstap bij kleine steekproeven wanneer je twee populatiegemiddeldes wil vergelijken
- $H_0: \sigma_1^2 = \sigma_2^2$ versus $H_a: \sigma_1^2 \leq \sigma_2^2$
- anders geschreven:

$$H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1 \qquad H_a: \frac{\sigma_1^2}{\sigma_2^2} \nleq 1$$

Kansdichtheid

aanpak?

$$\frac{(n_1 - 1)S_1^2}{\sigma_1^2} \sim \chi_{n_1 - 1}^2$$
$$\frac{(n_2 - 1)S_2^2}{\sigma_2^2} \sim \chi_{n_2 - 1}^2$$

indien populaties normaal verdeeld zijn

- nieuwe kansdichtheid:
 - \rightarrow quotiënt van twee χ^2 -verdeelde variabelen
 - → beide variabelen delen door aantal vrijheidsgraden
 - \rightarrow F-verdeelde kansvariabele

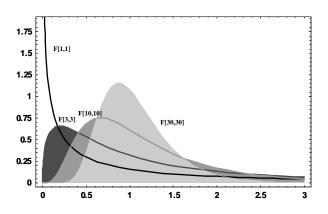
Fisher's *F*-verdeling

- 2 parameters:
 - $v_1 \rightarrow$ aantal vrijheidsgraden teller
 - $v_2 \rightarrow$ aantal vrijheidsgraden noemer

$$F_{v_1,v_2} = \frac{\chi_{v_1}^2/v_1}{\chi_{v_2}^2/v_2}$$

- ► R:
 - $\rightarrow P(F_{v_1,v_2} > f)$ "1-pf(f, v_1, v_2)"
 - \rightarrow kritieke waarden met "qf (1- α , v_1 , v_2)"
- Matlab:
 - $\rightarrow P(F_{v_1,v_2} > f)$ "1-fcdf(f, v_1, v_2)"
 - \rightarrow kritieke waarden met "finv $(1-\alpha, v_1, v_2)$ "

F-verdeling grafisch



Toetsingsgrootheid

$$\frac{\frac{(n_1-1)S_1^2}{\sigma_1^2}/(n_1-1)}{\frac{(n_2-1)S_2^2}{\sigma_2^2}/(n_2-1)} \sim F_{n_1-1,n_2-1}$$

vereenvoudigen levert

$$\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \sim F_{n_1 - 1, n_2 - 1}$$

► als nulhypothese H_0 : $\sigma_1^2 = \sigma_2^2$ waar is, dan

$$\frac{S_1^2}{S_2^2} \sim F_{n_1 - 1, n_2 - 1}$$

Rechtseenzijdige toets

- $H_a: \sigma_1^2 > \sigma_2^2$ of $H_a: \frac{\sigma_1^2}{\sigma_2^2} > 1$
- $ightharpoonup H_a$ aannemelijk indien

$$s_1^2 \gg s_2^2 \text{ of } \frac{s_1^2}{s_2^2} \gg 1$$

- H_a aanvaarden als $\frac{s_1^2}{s_2^2}$ groot is
 - \rightarrow groter dan kritieke waarde F_{α,n_1-1,n_2-1}
- ► *p*-waarde

$$\rightarrow p = P(F_{n_1 - 1, n_2 - 1} > \frac{s_1^2}{s_2^2}) = P(F_{n_1 - 1, n_2 - 1} > f)$$

Linkseenzijdige toets

- $H_a: \sigma_1^2 < \sigma_2^2$ of $H_a: \frac{\sigma_1^2}{\sigma_2^2} < 1$
- $ightharpoonup H_a$ aanvaarden indien

$$s_1^2 \ll s_2^2 \text{ of } \frac{s_1^2}{s_2^2} \ll 1$$

- H_a aanvaarden als $\frac{s_1^2}{s_2^2}$ klein is
 - ightarrow kleiner dan kritieke waarde $F_{1-\alpha,n_1-1,n_2-1}$
- ▶ *p*-waarde

$$\rightarrow p = P(F_{n_1 - 1, n_2 - 1} < \frac{s_1^2}{s_2^2}) = P(F_{n_1 - 1, n_2 - 1} < f)$$

Tweezijdige toets

$$H_a: \sigma_1^2 \neq \sigma_2^2$$
 of $H_a: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$

- ► H_a : $\sigma_1^2 \neq \sigma_2^2$ of H_a : $\frac{\sigma_1^2}{\sigma_2^2} \neq 1$ ► H_a aanvaarden indien $s_1^2 \ll s_2^2$ of $s_1^2 \gg s_2^2$ ► H_a aanvaarden indien $\frac{s_1^2}{s_2^2}$ te klein of te groot is
 - → kleiner dan linkerkritieke waarde

$$F_{1-\frac{\alpha}{2},n_1-1,n_2-1}$$

of groter dan rechterkritieke waarde

$$s_1^2 < s_2^2: p = 2P(F_{n_1-1,n_2-1} < \frac{s_1^2}{s_2^2})$$

$$s_1^2 > s_2^2: p = 2P(F_{n_1-1,n_2-1} > \frac{s_1^2}{s_2^2})$$

Voorbeeld

- $H_0: \sigma_1^2 = \sigma_2^2$ en $H_a: \sigma_1^2 \neq \sigma_2^2$
- ► steekproef 1: $s_1^2 = 0.1848$, $n_1 = 20$ steekproef 2: $s_2^2 = 0.2837$, $n_2 = 20$

$$f = \frac{s_1^2}{s_2^2} = \frac{0.1848}{0.2837} = 0.6514$$

kritieke waarde:

$$F_{1-\alpha/2,n_1-1,n_2-1} = F_{0.975,19,19} \approx 0.3958$$

- $f < F_{0.975,19,19} \Rightarrow \text{aanvaard } H_0$
- p-waarde = $2p(F_{19,19} < f) = 0.3583$

Twee populatieproporties

►
$$H_0: \pi_1 = \pi_2 \text{ versus } H_a: \pi_1 \underset{\neq}{\leq} \pi_2$$

► $\hat{P}_1 \sim N(\pi_1, \frac{\pi_1(1-\pi_1)}{n_1})$ (indien $n_1\pi_1 \ge 5$, $n_1(1-\pi_1) \ge 5$)

• $\hat{P}_2 \sim N(\pi_2, \frac{\pi_2(1-\pi_2)}{n_2})$ (indien $n_2\pi_2 \ge 5$, $n_2(1-\pi_2) \ge 5$)

• $\hat{P}_1 - \hat{P}_2 \sim N(\pi_1 - \pi_2, \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2})$

• als H_0 waar: $\pi_1 = \pi_2 = \pi$

• $N(0, \frac{\pi(1-\pi)}{n_1} + \frac{\pi(1-\pi)}{n_2})$

Toetsingsgrootheid

$$Z = \frac{\hat{P}_1 - \hat{P}_2 - 0}{\sqrt{\frac{\pi(1-\pi)}{n_1} + \frac{\pi(1-\pi)}{n_2}}} \sim N(0, 1)$$

- $\pi \text{ schatten m.b.v. } \overline{P} = \frac{n_1 \hat{P}_1 + n_2 \hat{P}_2}{n_1 + n_2}$
- beslissingsregel 1 (rechtseenzijdig):
 - $z > z_{\alpha}$: verwerp H_0
 - $z \le z_{\alpha}$: aanvaard H_0
- beslissingsregel 2 (rechtseenzijdig):

$$p = P(Z > z)$$

 $p < \alpha$: verwerp H_0

 $p \ge \alpha$: aanvaard H_0

Twee populatielocaties - onafhankelijke steekproeven

- Wilcoxon rangsomtoets
- ordinale gegevens
- ► H_0 : locatie 1 = locatie 2 H_a : locatie 1 \leq locatie 2
- vaak ook voor niet-normaal verdeelde kwantitatieve gegevens

Voorbeeld

- het pesticide DDT veroorzaakt stuiptrekkingen bij mensen en zoogdieren
- hoe worden deze stuiptrekkingen veroorzaakt?
- proefdieren: 6 met DDT vergiftigde ratten en 6 niet-vergiftigde ratten
- stimulatie van een zenuw in de poot van de rat
- elektrische reactie vertoont een scherpe piek gevolgd door een tweede veel kleinere piek
- amplitude van de tweede piek wordt gemeten als percentage van de eerste piek

Voorbeeld

DDT-groep	controlegroep
12.207	11.074
16.869	9.686
25.050	12.064
22.429	9.351
8.456	8.182
20.589	6.642

 $\begin{cases} H_0: \text{ geen verschil} \\ H_a: \text{ amplitude groter bij DDT vergiftigde ratten} \end{cases}$

 $\begin{cases} H_0 : \text{locatie } 1 = \text{locatie } 2 \\ H_a : \text{locatie } 1 > \text{locatie } 2 \end{cases}$

Werkwijze

- 1. kleinste steekproef → nummer 1
- 2. rangschik alle data van klein naar groot
- 3. geef rangnummers
- **4.** T_1 = som rangnummers steekproef 1
- 5. T_2 = som rangnummers steekproef 2
- 6. Als H_a waar zou zijn, dan zou T_1 te groot zijn: verwerp H_0 en aanvaard H_a indien

 T_1 > rechterkritieke waarde T_U

(cursus: appendix B)

Dl	DT-groep	controlegroep				
Meting	Rangnummer	Meting	Rangnummer			
12.207	8	11.074	6			
16.869	9	9.686	5			
25.050	12	12.064	7			
22.429	11	9.351	4			
8.456	3	8.182	2			
20.589	10	6.642	1			

$$t_1 = 8 + 9 + 12 + 11 + 3 + 10 = 53$$

$$n_1 = 6, n_2 = 6 \Rightarrow T_U = 50$$

a. $\alpha = 0.025$ eenzijdig; $\alpha = 0.05$ tweezijdig

n_1	3		3 4		5		6		1	7	8			9	10	
	TL	Tu	TL	Tu	Tı	Tu	TL	Tu	TL	$T_{\rm U}$	$T_{\rm L}$	Tu	TL	Tu	TL	Tu
3	5	16	6	18	6	21	.7	23	7	26	8	28	8	31	9	33
4	6	18	11	25	12	28	12	32	13	35	14	38	15	41	16	44
5	6	21	12	28	18	37	19	41	20	45	21	49	22	53	24	56
6	7	23	12	32	19	41	26	52	28	56	29	61	31	65	32	70
7	7	26	13	35	20	45	28	56	37	68	39	73	41	78	43	83
8	8	28	14	38	21	49	29	61	39	73	49	87	51	93	54	98
9	8	31	15	41	22	53	31	65	41	78	51	93	63	108	66	114
10	9	33	16	44	24	56	32	70	43	83	54	98	66	114	79	131

b. $\alpha = 0.05$ eenzijdig; $\alpha = 0.10$ tweezijdig

n ₁	3			4		5		5	7	7		8		9	- 1	0
	TL	Tu	TL	Tu	TL	Tu	TL	Tu	T_{L}	Tu	TL	Tu	$T_{\rm L}$	Tu	$T_{\rm L}$	Tu
3	6	15	7	17	7	20	.8	22	9	24	9	27	10	29	11	31
4	7	17	12	24	13	27	14	30	15	33	16	36	17	39	18	42
5	7	20	13	27	19	36	20	40	22	43	24	46	25	50	26	.54
6	8	22	14	30	20	40	28	50	30	54	32	58	33	63	35	- 67
7	9	24	15	33	22	43	30	54	39	66	41	71	43	76	46	80
8	9	27	16	36	24	46	32	58	41	71	52	84	54	90	57	95
9	10	29	17	39	25	50	33	63	43	76	54	90	66	105	69	111
10	11	31	18	42	26	54	35	67	46	80	57	95	69	111	83	127

D	DT-groep	controlegroep				
Meting	Rangnummer	Meting	Rangnummer			
12.207	8	11.074	6			
16.869	9	9.686	5			
25.050	12	12.064	7			
22.429	11	9.351	4			
8.456	3	8.182	2			
20.589	10	6.642	1			

- $t_1 = 8 + 9 + 12 + 11 + 3 + 10 = 53$
- $n_1 = 6, n_2 = 6 \Rightarrow T_U = 50$
- $t_1 > T_U$: verwerp nulhypothese
- conclusie: amplitude tweede piek groter bij DDT vergiftigde ratten

- tussen punten op de huid bestaat een natuurlijk elektrisch potentiaalverschil
- bevordert deze natuurlijke elektrische veldsterkte de genezing van wonden het beste?
- indien ja, dan vertraagt de genezing bij het wijzigen van de veldsterkte
- proefdieren : verdoofde watersalamanders

- sneetjes in beide achterpoten
- ► 1 achterpoot = experimenteel lidmaat (met elektrode om elektrisch veld te reduceren)
- 1 achterpoot = controlelidmaat (natuurlijke genezing)
- meet de snelheid waarmee de wonden genezen (in micrometers per uur)

Salamander	Experimenteel lidmaat	Controle lidmaat
1	24	25
2	23	13
3	47	44
4	42	45
5	26	57
6	46	42
7	38	50
8	33	36
9	28	35
10	28	38
11	21	43
12	27	31
13	25	26
14	45	48

Wilcoxon rangtekentoets

```
H_0: geen verschil in genezingstempo H_a: lager bij gewijzigde veldsterkte
```

```
\begin{cases} H_0 : \text{locatie } 1 = \text{locatie } 2 \\ H_a : \text{locatie } 1 < \text{locatie } 2 \end{cases}
```

Werkwijze

- 1. verschilvariabele
- 2. rangschik absolute waardes
- 3. geef rangnummers + of teken

4. tel
$$\begin{cases} \text{positieve} \\ |\text{negatieve}| \end{cases}$$
 rangnummers op:
$$\begin{cases} T_{+} \\ T_{-} \end{cases}$$

5. als H_0 waar: $T_+ \approx T_-$

als H_a waar: $\begin{cases} \text{veel negatieve verschillen} \\ \text{weinig positieve verschillen} \end{cases}$

 $\Leftrightarrow \begin{cases} T_+ \text{klein} \\ T_- \text{groot} \end{cases} \to T_+ \text{ vgln. met linkerkritieke}$

waarde

(cursus: appendix C)

Salamander	Experimenteel lidmaat	Controle lidmaat	Verschil	Verschil	Rangnummers Verschil	Rangnummers met teken
1	24	25	-1	1	1.5	-1.5
2	23	13	10	10	10.5	10.5
3	47	44	3	3	4.5	4.5
4	42	45	-3	3	4.5	-4.5
5	26	57	-31	31	14	-14
6	46	42	4	4	7.5	7.5
7	38	50	-12	12	12	-12
8	33	36	-3	3	4.5	-4.5
9	28	35	-7	7	9	-9
10	28	38	-10	10	10.5	-10.5
11	21	43	-22	22	13	-13
12	27	31	-4	4	7.5	-7.5
13	25	26	-1	1	1.5	-1.5
14	45	48	-3	3	4.5	-4.5

Voorbeeld

hypothesen

```
\begin{cases} H_0: \text{ geen verschil in genezingstempo} \\ H_a: \text{ lager bij gewijzigde veldsterkte} \end{cases}
\begin{cases} H_0: \text{ locatie } 1 = \text{ locatie } 2 \\ H_a: \text{ locatie } 1 < \text{ locatie } 2 \end{cases}
```

$$n = 14 \Rightarrow T_0 = 26$$

Eenzijdig	Tweezijdig	n = 5	n = 6	n = 7	n = 8	n = 9	n = 10
$\alpha = 0.05$	$\alpha = 0.10$	1	2	4	6	8	11
$\alpha = 0.025$	$\alpha = 0.05$		1	2	4	6	8
$\alpha = 0.01$	$\alpha = 0.02$		III DE	0	2	3	5
$\alpha = 0.005$	$\alpha = 0.01$	11118			0	2	3
	100	n = 11	n = 12	n = 13	n = 14	n = 15	n = 16
$\alpha = 0.05$	$\alpha = 0.10$	14	17	21	26	30	36
$\alpha = 0.025$	$\alpha = 0.05$	11	14	17	21	25	30
$\alpha = 0.01$	$\alpha = 0.02$	7	10	13	16	20	24
$\alpha = 0.005$	$\alpha = 0.01$	5	7	10	13	16	19
		n = 17	n = 18	n = 19	n = 20	n = 21	n = 22
$\alpha = 0.05$	$\alpha = 0.10$	41	47	54	60	68	75
$\alpha = 0,025$	$\alpha = 0.05$	35	40	46	52	59	66
$\alpha = 0.01$	$\alpha = 0.02$	28	33	38	43	49	56
$\alpha = 0.005$	$\alpha = 0.01$	23	28	32	37	43	49
		n = 23	n = 24	n = 25	n = 26	n = 27	n = 28
$\alpha = 0.05$	$\alpha = 0.10$	83	92	101	110	120	130
$\alpha = 0.025$	$\alpha = 0.05$	73	81	90	98	107	117
$\alpha = 0.01$	$\alpha = 0.02$	62	69	77	85	93	102
$\alpha = 0,005$	$\alpha = 0.01$	55	61	68	76	84	92
		n = 29	n = 30	n = 31	n = 32	n = 33	n = 34
$\alpha = 0.05$	$\alpha = 0.10$	141	152	163	175	188	201
$\alpha = 0.025$	$\alpha = 0.05$	127	137	148	159	171	183
$\alpha = 0.01$	$\alpha = 0.02$	111	120	130	141	151	162
$\alpha = 0.005$	$\alpha = 0.01$	100	109	118	120	-	

Voorbeeld

hypothesen

```
\begin{cases} H_0: \text{ geen verschil in genezingstempo} \\ H_a: \text{ lager bij gewijzigde veldsterkte} \\ H_0: \text{ locatie } 1 = \text{ locatie } 2 \\ H_a: \text{ locatie } 1 < \text{ locatie } 2 \end{cases}
```

- ► $t_+ = 22.5$
- $n = 14 \Rightarrow T_0 = 26$
- $t_+ < T_0$: verwerp H_0
- conclusie: gewijzigde veldsterkte vertraagt het genezingsproces