

# Testiranje razlika između proporcija



**PREDAVANJE BR. 2**

# Standardna greška proporcije



- Podaci koji nijesu rezultat kvantitativnih mjerenja
- Frekvencije pojavljivanja – proporcije
- Primjer: da li su neki ljudi oboljeli od neke bolesti ili ne?
- Proporcija uzorka ima svoju standardnu grešku - standardna greška proporcije
- $p$  – relativna učestalost pojavljivanja nekog svojstva

$$SD_p = \sqrt{\frac{pq}{N}}$$

# Razlika između proporcija



Značajnost razlika između drugih stat.mjera utvrđuje se postupcima koji su isti kao kod razlika između aritmetičkih sredina.

1. Velikih nezavisnih uzoraka
2. Velikih zavisnih uzoraka
3. Malih nezavisnih uzoraka
4. Malih zavisnih uzoraka

# Razlike između proporcija kod velikih nezavisnih uzoraka



$$H_0 : \pi_1 = \pi_2$$

$$H_1 : \pi_1 \neq \pi_2$$

$$SD_{p_1-p_2} = \sqrt{SD_{p_1}^2 + SD_{p_2}^2} = \sqrt{\frac{p_1 q_1}{N_1} + \frac{p_2 q_2}{N_2}}$$

$$t = \frac{p_1 - p_2}{SD_{p_1-p_2}}$$

Ako su uzorci veliki, a  $p$  i  $q$  nema ekstremno malu ili veliku vrijednost, može se pretpostaviti normalna raspodjela. Ili, ako je  $Np$  i  $Nq$  kod oba uzorka veće od 5 (po drugima, veće od 10).

# Razlike između AS velikih zavisnih uzoraka



$$H_0 : \pi_1 = \pi_2$$

$$H_1 : \pi_1 \neq \pi_2$$

$$SD_{p_1-p_2} = \sqrt{SD_{p_1}^2 + SD_{p_2}^2 - 2r_{1,2}SD_{p_1}SD_{p_2}}$$

Postoji i jednostavniji postupak gdje možemo izbjeći računanje korelacije, ukoliko poznajemo sve rezultate ispitanika i njihove individualne karakteristike koje možemo tabelarno prikazati.

# Razlike između AS malih nezavisnih uzoraka



- Ako su  $N_1$  i  $N_2$  mali (manji od 100) i ako su proporcije ekstremne (veće od 0.9 ili manje od 0.1)
- Računa se kombinovana proporcija

$$p = \frac{p_1 N_1 + p_2 N_2}{N_1 + N_2}$$

$$SD_{p_1 - p_2} = \sqrt{pq \left( \frac{1}{N_1} + \frac{1}{N_2} \right)}$$

# Razlike između AS malih zavisnih uzoraka



- Kao i kod velikih zavisnih, treba unijeti rezultate u tablicu

$$SD_{p_1 - p_2} = \sqrt{\frac{d + a}{N}}$$

- Primjer: Ispitanici su podvgnuti ispitivanju psihomotorike pomoću dva testa koji daju rezultate “zadovoljava” i “nezadovoljava”. Je li razlika stat. značajna, tj. da li je drugi test lakši od prvog?

		Drugi test		
		Nijesu zadovoljili	Zadovoljili	Ukupno
Prvi test	Zadovoljili	5 (0,05) <sup>a</sup>	55 (0,55) <sup>b</sup>	60 (0,60) <sup>p<sub>1</sub></sup>
	Nijesu zadovoljili	25 (0,25) <sup>c</sup>	15 (0,15) <sup>d</sup>	40 (0,40)
	Ukupno	30 (0,30)	70 (0,7) <sup>p<sub>2</sub></sup>	100 (1,00)

# Primjeri



## Primjer 1:

U grupi od 230 muškaraca, 90 su oboljeli od neke bolesti, a u grupi od 260 žena, 170 ih je oboljelo od te iste bolesti. Postoji li značajna razlika između muškaraca i žena u frekvenciji obolijevanje od te bolesti?

**Pol \* Bolest Crosstabulation**

Count

		Bolest		Total
		Da	Ne	
Pol	M	90	140	230
	Ž	170	90	260
Total		260	230	490

**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
Bolest_nova	Equal variances assumed	4.102	.043	6.011	488	.000	.263	.044	.177	.348
	Equal variances not assumed			6.001	477.455	.000	.263	.044	.177	.349



# Primjeri



## Primjer 1:

U grupi od 230 muškaraca, 90 su oboljeli od neke bolesti, a u grupi od 260 žena, 170 ih je oboljelo od te iste bolesti. Postoji li značajna razlika između muškaraca i žena u frekvenciji obolijevanje od te bolesti?

- $p_1 = \frac{90}{230} = 0,391$
- $p_2 = \frac{170}{260} = 0,654$
- $p_2 - p_1 = 0,263$
- $N_1$  i  $N_2 > 100$
- $SD_{p_1-p_2} = \sqrt{\frac{p_1q_1}{N_1} + \frac{p_2q_2}{N_2}} = \sqrt{0,0019} = 0,0436$
- $t = \frac{0,263}{0,0436} = 6,02$

# Primjeri



## Primjer 2

U grupi od 50 dječaka, 24 su visine preko 122 cm, dok su u skupini od 60 djevojčica njih 18 viši od 122 cm. Može li se ta razlika u proporciji visoke djece smatrati statistički značajnom uz rizik greške 5%?

**Visina \* Pol Crosstabulation**

Count		Pol		Total
		Dječak	Djevojčica	
Visina	preko 122cm	24	18	42
	ispod 122 cm	26	42	68
Total		50	60	110

**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
Visina	Equal variances assumed	9.075	.003	1.951	108	.054	.180	.092	-.003	.363
	Equal variances not assumed			1.935	100.606	.056	.180	.093	-.005	.365

# Primjeri



## Primjer 2

U grupi od 50 dječaka, 24 su visine 122 cm, dok su u skupini od 60 djevojčica njih 18 viši od 122 cm. Može li se ta razlika u proporciji visoke djece smatrati statistički značajnom uz rizik greške 5%?

$$N_1 \text{ i } N_2 < 100$$

$$p_1 = \frac{24}{50} = 0,48$$

$$p_2 = \frac{18}{60} = 0,3$$

$$N_1=50 \quad N_2=60$$

$$p = \frac{p_1 N_1 + p_2 N_2}{N_1 + N_2} = 0,38$$
$$q = 0,62$$

$$SD_{p_1-p_2} = \sqrt{pq\left(\frac{1}{N_1} + \frac{1}{N_2}\right)} = 0,093$$

$$t = \frac{0,18}{0,093} = 1,94$$

# Primjeri



## Primjer 3

100 ispitanika podvrnuto je ispitivanju psihomotorike pomoću dva testa koji su takve prirode da daju samo alternativne rezultate: “zadovoljava” i “ne zadovoljava”. U prvom testu 60 ispitanika zadovoljava, a u drugom 70. Je li ta razlika statistički značajna, tj. je li drugi test stvarno lakši od prvog?

		test2		Total
		Zadovoljava	Nezadovoljava	
test1	Zadovoljava	55	5	60
	Nezadovoljava	15	25	40
Total		70	30	100

			test2		Total
			Zadovoljava	Nezadovoljava	
test1	Zadovoljava	Count	55	5	60
		% of Total	55.0%	5.0%	60.0%
	Nezadovoljava	Count	15	25	40
	% of Total	15.0%	25.0%	40.0%	
Total	Count	70	30	100	
	% of Total	70.0%	30.0%	100.0%	

# Primjeri



## Primjer 3

100 ispitanika podvrnuto je ispitivanju psihomotorike pomoću dva testa koji su takve prirode da daju samo alternativne rezultate: “zadovoljava” i “ne zadovoljava”. U prvom testu 60 ispitanika zadovoljava, a u drugom 70. Je li ta razlika statistički značajna, tj. je li drugi test stvarno lakši od prvog?

**Paired Samples Test**

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	test1 - test2	.100	.438	.044	.013	.187	2.283	99	.025

# Primjeri



## Primjer 3

100 ispitanika podvrnuto je ispitivanju psihomotorike pomoću dva testa koji su takve prirode da daju samo alternativne rezultate: “zadovoljava” i “ne zadovoljava”. U prvom testu 60 ispitanika zadovoljava, a u drugom 70. Je li ta razlika statistički značajna, tj. je li drugi test stvarno lakši od prvog?

$$SD_{p_1-p_2} = \sqrt{\frac{d+a}{N}} = 0,0447$$

$$d = 0,15$$

$$a = 0,05$$

$$N = 100$$

$$t = \frac{0,10}{0,0447} = 2,24$$

# Primjeri



## Primjer 4

Na reprezentativnim uzorcima studenata (660) i studentkinja (360) UCG, utvrđeno je da 57%, odnosno 38% čita studentski list. Da li je razlika između procenata statistički značajna, tj. da li značajno veći broj studenata čita studentski list?

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
Čita	Equal variances assumed	10.515	.001	-5.865	1018	.000	-.189	.032	-.252	-.126
	Equal variances not assumed			-5.897	749.867	.000	-.189	.032	-.252	-.126

# Provjera znanja



1. 30 ispitanika je rješavalo test na osnovu koga su, izjednačavanjem po parovima, formirane dvije grupe od po 15 ispitanika, eksperimentalna i kontrolna. Obije grupe su učile isto gradivo ali različitim metodama. Na kraju eksperimenta, testom je izmjereno znanje gradiva koje su ispitanici učili. Rezultati testa su dati u sljedećoj tabeli:

<b>E</b>	98	85	103	90	79	84	100	86	81	112	76	94	97	82	88
<b>K</b>	86	77	93	78	67	70	94	80	68	102	63	76	85	69	78

Utvrđiti da li različite metode učenja daju različite efekte u učenju određenog gradiva.



# Provjera znanja



1. 30 ispitanika je rješavalo test na osnovu koga su, izjednačavanjem po parovima, formirane dvije grupe od po 15 ispitanika, eksperimentalna i kontrolna. Obije grupe su učile isto gradivo ali različitim metodama. Na kraju eksperimenta, testom je izmjereno znanje gradiva koje su ispitanici učili. Rezultati testa su dati u sljedećoj tabeli:

$$t = \frac{M_D}{SD_{MD}} = \frac{M_D}{\sqrt{\frac{\sum d^2}{N(N-1)}}} = \frac{11,26667}{\sqrt{\frac{134,933}{15 * 14}}} = \frac{11,26667}{\sqrt{0,642538}} = 14,055$$

Utvrđiti da li različite metode učenja daju različite efekte u učenju određenog gradiva.

# Provjera znanja



2. Jedan test ličnosti zadat je trinaestorici učenika, od koji su se, u školi, osmorica ponašala neupadljivo, a petorica ekcesno. Dobijene su sljedeće mjere:

<b>G1</b>	110	112	95	105	111	97	112	102
<b>G2</b>	115	112	109	112	117			

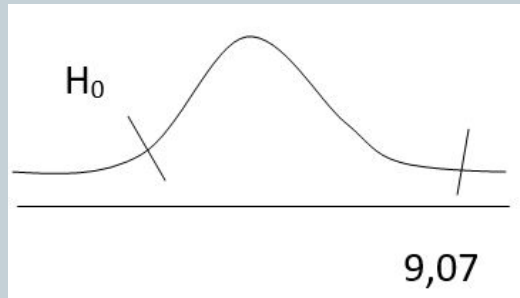
Da li se ove grupe učenika značajno razlikuju po prosječnim rezultatima na testu ličnosti?

# Provjera znanja - nastavak



$$F = \frac{\text{veća } SD^2}{\text{manja } SD^2} = \frac{47,143}{9,5} = 4,962406$$

$$F_{\alpha, n_1-1, n_2-1} = F_{0,05, 7, 4} = 9,07$$



$$Zaj. SD = \sqrt{\frac{SD_1^2(N_1 - 1) + SD_2^2(N_2 - 1)}{(N_1 - 1) + (N_2 - 1)}} = 5,78399$$

$$SD_{M_1 - M_2} = Zaj. SD \sqrt{\frac{N_1 + N_2}{N_1 * N_2}} = 3,297382$$

$$t = \frac{M_1 - M_2}{SD_{M_1 - M_2}} = \frac{105,5 - 113}{3,297382} = -2,27453$$

# Provjera znanja



3. Na testu verbalnih sposobnosti, grupa ispitanika čiji roditelji imaju visoko ili više obrazovanje i grupa ispitanika čiji roditelji imaju osnovno ili srednje obrazovanje, postigle su sljedeće rezultate:

Obrazovanje roditelja			
		Osnovno/srednje	Visoko/Više
Rezultati	M	16,92	17,90
	SD	4,94	4,21
	N	48	56

Utvrđiti da li ove dvije grupe ispitanika imaju i značajno različit prosječan uspjeh na primijenjenom testu.

# Provjera znanja



3. Na testu verbalnih sposobnosti, grupa ispitanika čiji roditelji imaju visoko ili više obrazovanje i grupa ispitanika čiji roditelji imaju osnovno ili srednje obrazovanje. Utvrditi da li ove dvije grupe ispitanika imaju i značajno različit prosječan uspjeh na primijenjenom testu.

- $SD_{M_1-M_2} = \sqrt{\frac{SD_1^2}{N_1} + \frac{SD_2^2}{N_2}}$
- $t = \frac{M_1 - M_2}{SD_{M_1-M_2}} = \frac{16,92 - 17,9}{0,90824} = -1,079$

# Provjera znanja



4. Od 100 devetogodišnje djece, zadatak iz jednog testa njih 72 rješava tačno, a od 100 desetogodišnjaka – 78. Da li je ovaj zadatak značajno teži za djecu od devet godina?

- $N_1=100$     $p_1=0,72$     $q_1=0,28$
- $N_2=100$     $p_2=0,78$     $q_2=0,22$
- $SD_{p_1-p_2} = \sqrt{\frac{p_1q_1}{N_1} + \frac{p_2q_2}{N_2}} = 0,06109$
- $t = \frac{0,06}{0,06109} = 0,982156$