

UPPGIFT 1

- Vid slumpmässigt urval har varje individ en känd sannolikhet att komma med i urvalet
- Resultatet går att generalisera till populationen när man gjort slumpmässigt urval

UPPGIFT 2

A) Konstanttermen: Den som inte får någon utbildning alls presterar c:a 50 (gäller om $x=0$ observerats)

För varje ytterligare utbildningstimma ökar prestationen med i genomsnitt 5.

B) Utbildningstiden bör vara 6 så att $50+5*6=80$

UPPGIFT 3

A) $p=142/327=0,43$

95%-igt konfidensintervall för andel kalsongbärare i populationen:

$$p \pm 1,96 \sqrt{\frac{p(1-p)}{n}}$$

Förutsättningar: OSU, oberoende mätningar, $np(1-p) > 9$

$np(1-p)=327*0,43*0,57=80$; tillräckligt stort stickprov!

$$p \pm 1,96 \sqrt{\frac{p(1-p)}{n}} = 0,43 \pm 1,96 \sqrt{\frac{0,43(1-0,43)}{327}} = 0,43 \pm 0,05$$

Nedre gräns= $0,43-0,05=0,38$

Övre gräns= $0,43+0,05=0,48$

Slutsats: Med 95% säkerhet finns andelen kalsongbärare i intervallet 38-48%

B) För halverat konfidensintervall krävs ett fyra gånger så stort stickprov, dvs $4*327=1308$

Optimal lösning: Räkna med det mest pessimistiska π utifrån pilotundersökningen (0,48) mha formel på sid 13 i formelsamlingen:

$$n > \left(\frac{z}{m} \right)^2 p^*(1-p^*) = \left(\frac{1,96}{0,025} \right)^2 * 0,48 * 0,52 = 1534$$

UPPGIFT 4

Person nr	Genomsnittligt värde före försöket	Genomsnittligt värde under försöket	d	d ²
1	12	9	3	9
2	14	11	3	9
3	8	9	-1	1
4	11	10	1	1
5	10	10	0	0
6	15	17	-2	4
7	13	10	3	9
8	14	11	3	9
			10	42

Medelvärde= 1,25

$$s = \sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n-1}} = \sqrt{\frac{42 - \frac{10^2}{8}}{7}} = 2,05$$

H₀: GI-mat påverkar inte blodsockerhalten (D=0)

H₁: GI-mat sänker blodsockerhalten (D>0)

Förutsättningar för t-test: OSU, oberoende mätningar, variabeln är approximativt normalfördelad

n-1=7, vilket medför att t-fördelningen med 7 frihetsgrader ska användas

H₀ förkastas om t>1,895 eftersom testet genomförs med 5% signifikansnivå

$$t = \frac{d_{obs} - D_0}{s_d / \sqrt{n}} = \frac{1,25 - 0}{2,05 / \sqrt{8}} = 1,72 < 1,895$$

Resultatet är icke signifikant. H₀ accepteras.

Group Statistics

Mätning	N	Mean	Std. Deviation	Std. Error Mean
GI före	8	12,13	2,357	,833
GI efter	8	10,88	2,588	,915

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
GI	Equal variances assumed	,119	,736	1,010	14	,330	1,250	1,237	-1,404	3,904
	Equal variances not assumed			1,010	13,879	,330	1,250	1,237	-1,406	3,906

B) Styrka=sannolikheten att förkasta H_0 när den är falsk

Förkastelsegräns: H_0 förkastas när $t > 1,895$ dvs om

$$d_{\text{obs}} > 0 + 1,895 \text{ medelfel} = 1,895 * \frac{2,05}{\sqrt{8}} = 1,37$$

$$\text{Sökt} = \Pr(d > 1,37 | D = 2,5) = \Pr\left(t > \frac{1,37 - 2,5}{\frac{2,05}{\sqrt{8}}}\right) = \Pr(t > -1,56) > 0,90 \text{ eftersom sannolikheten att } t < -$$

1,415 är 10% enligt t-tabellen (7 df)

C) Teckentest/mediantest, där differenserna testas m h a binomialfördelningen.

UPPGIFT 5

A)

A = Händelsen att ett misstag sker på måndag

B = Händelsen att ett misstag sker den sista timmen

$$P(A) = 0,25$$

$$P(B) = 0,20$$

$$P(A \text{ och } B) = 0,04$$

Det är enklast att upprätta en sannolikhetsstabell och placera in de kända sannolikheterna och därfter beräkna de som saknas.

	A	Icke-A	
B	0,04	0,16	0,20
Icke-B	0,21	0,59	0,8
	0,25	0,75	1,0

Sökt: $P(\text{icke-B} | A)$

$$P(A | B) = \frac{P(A \text{ och } B)}{P(B)} \Rightarrow P(\bar{B} | A) = \frac{P(A \text{ och } \bar{B})}{P(A)} = \frac{0,21}{0,25} = 0,84$$

B)

Nej, eftersom vid oberoende gäller $P(A) \cdot P(B) = P(A \text{ och } B)$

$$0,25 \cdot 0,20 \neq 0,04$$

Man kan också jämföra sannolikheterna $P(\bar{B} | A)$ och $P(\bar{B})$ vilka inte är lika. Jämför sannolikheterna 0,84 med 0,8 (1-0,2)

Händelserna A och B är beroende!

C)

X = Antal misstag under måndagens sista timme

n = 100 misstag

p = Sannolikheten att ett misstag inträffar under måndagens sista timme = 0,04

Högst ett misstag får inträffa innebär att de eftersökta värdena på X är noll och ett.

X följer en binomialfördelning med n=100 och p = 0,04

$$P(X \leq 1 | X \text{ är Bi}(n=100; p=0,04)) = P(X=0) + P(X=1)$$

$$pr(X = x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$$

$$pr(X = 0) = \binom{100}{0} \cdot 0,04^0 \cdot (1-0,04)^{100-0} = 0,0169$$

$$pr(X = 1) = \binom{100}{1} \cdot 0,04^1 \cdot (1-0,04)^{100-1} = 0,0703$$

Sannolikheten att högst ett misstag inträffar under måndagens sista timme är 0,0872

D)

Om frågan var kopplad till den årliga kostnaden som utgörs av personrelaterade misstag under måndagens sista timme så löser man den enligt följande:

X = Antal misstag under måndagens sista timme

n = 100 misstag

p = Sannolikheten att ett misstag inträffar under måndagens sista timme = 0,04

$$E(X) = n \cdot p = 100 \cdot 0,04 = 4$$

$$\sigma(X) = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{100 \cdot 0,04 \cdot 1 - 0,04} = 1,96$$

Man frågar efter förväntat värde och standardavvikelse för kostnaden under året.

Använd linjär transformation:

Y = Kostnaden för hela året som sker under måndagens sista timme.

$$Y = 46 \text{ (veckor)} \cdot 800 \text{ (kr/misstag)} \cdot X = 36\,800 \cdot X$$

$$Y = a \cdot X + b$$

$$E(Y) = a \cdot E(X) + b$$

$$Var(Y) = a^2 \cdot Var(X) \Rightarrow \sigma_Y = |a| \cdot \sigma_X$$

I vårt fall $Y = 36\,800 \cdot X$, är a = 36 800 och b = 0

Väntevärdet för kostnaden:

$$E(Y) = 36\,800 \cdot 4 = 147\,200$$

Standardavvikelsen för kostnaden:

$$\sigma_Y = 36\,800 \cdot 1,96 = 72\,128$$

Uppgift 6

A, B)

- Ålder (år)

Kvantitativ, Kontinuerlig, Kvotskala

Beräkna medelvärde och standardavvikelse

Presentera histogram

Klassindelad frekvenstabell

- Kön (man/kvinna)

Kvalitativ, nominalskala

Presentera typvärde

Stapeldiagram, cirkeldiagram

Frekvenstabell

- Ersättningsnivå före rehabilitering (25 %, 50 %, 75 % och 100 %)

Kvantitativ, ordinalskala

Beräkna median

Stapeldiagram eller cirkeldiagram

Frekvenstabell

- Ersättningsnivå efter rehabilitering (25 %, 50 %, 75 % och 100 %)

Kvantitativ, ordinalskala

Beräkna median

Stapeldiagram eller cirkeldiagram

Frekvenstabell

- Tid från första sjukskrivningsdag till initierat rehabiliteringsärende (dagar)

Kvantitativ, diskret, kvotskala

Beräkna medelvärde och standardavvikelse

Histogram

Klassindelad frekvenstabell

- Tid från initierat rehabiliteringsärende till påbörjad rehabilitering (dagar)

Kvantitativ, diskret, kvotskala

Beräkna medelvärde och standardavvikelse

Histogram

Klassindelad frekvenstabell

- Rehabiliteringens längd (dagar)

Kvantitativ, diskret, kvotskala

Beräkna medelvärde och standardavvikelse

Histogram

Klassindelad frekvenstabell

C)

Man kan beräkna medelvärden för tid innan initierat rehabiliteringsärende för män respektive kvinnor och jämföra dessa. Man kan åskådliggöra dessa i ett stapeldiagram

D)

Man kan studera förändringar i ersättningsnivåerna. Sänkt, oförändrat och höjt. Gruppindela ålder och skapa ett antal klasser. Illustreras lämpligast med korstabell där ålder summeras till 100 %. Man kan jämföra andelen som sänkt sin ersättningsgrad mellan unga och gamla.