



Stockholms universitet

OBS! Läs noga igenom anvisningarna i tentamen, t.ex. hur du ska skriva svaren. Det är ditt ansvar som student att följa de anvisningar som ges.

NOTE! Read the examination instructions carefully, e.g. how to write the answers. It is your responsibility as a student to follow the given instructions.

Skriv din anonymiseringskod och dagens datum på allt material du lämnar in.
(Enter your anonymization code and today's date on all submitted materials)

Anonymiseringskod (Anonymization code)	3	1	1	-	0	0	0	4	-	L	C	B
Datum (Date YYYY-MM-DD)	2023-03-20							Plats nr. (Seat No.)	40			

Kurs/Kurskod (Course/Course code)	ST1101
Kursmoment (Course component)	SDAI Tentamen 2

Fylls i av tentamensvärd (To be filled in by invigilator)

Direkt i skrivning: (kryss)		Svarsblankett: (kryss)		Lösa svarsblad: (antal)	3
--------------------------------	--	---------------------------	--	----------------------------	---

Lämnat in blankt: (kryss)		Dator: (kryss)	
------------------------------	--	-------------------	--

Inlämningstid: 12 : 55 Signatur tentamensvärd: PUG

Fylls i av lärare/examinator (To be filled in by teacher/examinator)

Betyg:	A	Poäng:	96
--------	---	--------	----

Signatur rättande lärare/examinator:



Datum: (Date YYYY-MM-DD)	2023-03-20	Kurs/Kurskod: (Course/Course code)	ST1101	Sidnr.: (Page no.)									
Anonymiseringskod (Anonymization code)	3	1	1	-	0	0	0	4	-	L	C	B	1

1.	$P(A) = 0,3$ $P(B) = 0,2$	Uppg.nr.: (Task no.)	1
a)	$P(A \cap B) = P(B A) \cdot P(A)$ $P(B A) = P(B)$ eftersom A och B är oberoende. $P(A \cap B) = P(B) \cdot P(A)$ $P(A \cap B) = 0,2 \cdot 0,3$ $P(A \cap B) = 0,06$ Svar: Sannolikheten att både A och B inträffar är 0,06.	Lärarens kommentar: (Teacher's note)	
b)	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ $P(A \cup B) = 0,3 + 0,2 - 0,06$ $P(A \cup B) = 0,44$ Svar: Sannolikheten att åtminstone A eller B inträffar är 0,44.		a. 4 b. 4 c. 4 d. 4
c)	$P(A^c \cap B^c) = 1 - P(A \cup B)$ $P(A^c \cap B^c) = 1 - 0,44$ $P(A^c \cap B^c) = 0,56$ Svar: Sannolikheten att ingen av händelserna A och B inträffar är 0,56.		
d)	$P(A B) = P(A)$ eftersom A och B är oberoende. $P(A B) = 0,3$ Svar: Eftersom A och B är två oberoende händelser så påverkas inte sannolikheten för A av att vi vet att B har inträffat. Det innebär att sannolikheten för A givet att B har inträffat är samma som sannolikheten för att A inträffar, vilket är 0,3.		
		Poäng: (Points)	16

2. Låt A vara händelsen att jag kommer i tid till föreläsningen.

$P(A) = 0,8 = p$

$P(A^c) = 1 - 0,8 = 0,2 = q$

a) $P(\text{föreläsning på måndag och onsdag}) = P(A^c) \cdot P(A) \cdot P(A) \cdot P(A) = P(A^c) \cdot P(A) \cdot P(A) \cdot P(A)$

$= 0,2 \cdot 0,8 \cdot 0,2 \cdot 0,8 \cdot 0,8 \cdot 0,8 = 0,02048$

Svar: Sannolikheten att jag blir föreläsad enbart på måndag och onsdag är 0,02048.

b) $P(\text{första föreläsningen på fredag})$

Geometrisk fördelning $X \sim \text{Geo}(0,2)$ $X = \text{antal dagar till första föreläsningen}$

$P(X = x) = q^{x-1} \cdot p$
 $q = 0,8 \quad p = 0,2$
 $P(X = 5) = 0,8^{5-1} \cdot 0,2 = 0,08192$

Svar: Sannolikheten att veckans första föreläsning blir på fredag är 0,08192.

2a) 4
 2b) 4
 2c) 6
 2d) 6

c) $P(\text{föreläsning 3 av 5 dagar}) \quad n = 5 \quad p = 0,2$

Binomialfördelning $X \sim \text{Binom}(5; 0,2)$ $X = \text{antal föreläsningar under veckan}$

$P(X = x) = n \cdot C_x \cdot p^x \cdot q^{n-x}$
 $P(X = 3) = 5 \cdot C_3 \cdot 0,2^3 \cdot 0,8^{5-3} = 5! \cdot 0,2^3 \cdot 0,8^2$

$P(X = 3) = \frac{5!}{(5-3)!3!} \cdot 0,08 \cdot 0,64 = \frac{5! \cdot 0,08 \cdot 0,64}{2! \cdot 3!} = 0,00512$

$= \frac{2!}{5!} \cdot 0,00512 = 10 \cdot 0,00512 = 0,0512$

Svar: Sannolikheten att jag blir föreläsad 3 av 5 dagar är 0,0512.

d) $X \sim \text{Binom}(5; 0,2)$ $X = \text{antal föreläsningar under veckan}$

$E(X) = np$
 $E(X) = 5 \cdot 0,2 = 1$

Med förväntade antal föreläsningar är en gång per vecka.

$E(100X) = 100 \cdot E(X)$

$= 100 \cdot 1$

$= 100$

Svar: Min förväntade kostnad under veckan är 100 kr.



Datum: (Date YYYY-MM-DD)	2023-03-20	Kurs/Kurskod: (Course/Course code)	ST1101	Sidnr.: (Page no.)	
Anonymiseringskod (Anonymization code)	3 1 1 - 0 0 0 4 - L C B				3

3.2) p okänt
 $\hat{p} = \frac{172}{230} = 0,74783$

$$\hat{p} \overset{\text{approx}}{\sim} N\left(\hat{p}, \sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$$

Antaganden: $n\hat{p} \geq 10 \rightarrow 230 \cdot 0,74783 = 172 > 10$

$$n\hat{q} \geq 10 \rightarrow 230 \cdot (1 - 0,74783) = 58 > 10$$

$$n \geq 30 \rightarrow n = 230 > 30$$

rimligt att stickprovet $< 10\%$ av populationen $\rightarrow 230 < 10\%$ av populationen

p och q är okända. Använder mig därför av \hat{p} och \hat{q} .

$$\sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{0,74783 \cdot (1 - 0,74783)}{230}} = 0,02863$$

$$\hat{p} \overset{\text{approx}}{\sim} N(0,74783, 0,02863)$$

Den normalapproximerade samplingfördelningen för \hat{p} är centrerad kring $\hat{p} = 0,74783$

b) $\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \alpha = 1 - 0,95 = 0,05$

$$0,74783 \pm z_{0,025} \cdot \sqrt{\frac{0,74783 \cdot (1 - 0,74783)}{230}}$$

$z_{0,025}$ hittas i normalfördelningstabell för sannolikhet $1 - 0,025 = 0,975$
 $z_{0,025} = 1,96$

$$0,74783 \pm 1,96 \cdot \sqrt{\frac{0,18858}{230}}$$

$$0,74783 \pm 1,96 \cdot 0,028634$$

$$0,74783 \pm 0,056123$$

$$0,74783 - 0,056123 = 0,691703$$

$$0,74783 + 0,056123 = 0,803949$$

$$(0,6917, 0,8039)$$

95% av alla konfidensintervall som skapas med formeln $\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ ur stickprov från populationen kommer att innehålla den sanna populationsandelen av de som skulle rekommendera företagets service till andra personer.

Uppg.nr.:
(Task no.)

3

Lärarens kommentar:
(Teacher's note)

3a.) 10,5

3b.) 9

Poäng:
(Points)

4 a) $H_0: \mu = 54,3$ minuter
 $H_A: \mu \neq 54,3$ minuter
 Dubbelstidigt test

$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$
 $X = 66,8$
 $s = 22,26$
 $n = 5$

$T = \frac{66,8 - 54,3}{22,26/\sqrt{5}}$
 $T = \frac{12,5}{9,955}$
 $T = 1,255654$

$t_{crit} = t_{0,05/2, 5-1}$
 $t_{crit} = 2,776$
 För att kunna förkasta nollhypotesen ska $|T| > t_{crit}$

$1,255654 < 2,776$
 $|T| < t_{crit}$
 Kan inte förkasta H_0
 med 5% signifikansnivå

Utifrån hypotestestet kan vi inte förkasta H_0 att medeltid för deltagarna inte har förändrats. Trots att medelvärdet för de slumpmässigt utvalda kunderna efter bytet av spelisterna är högre än den genomsnittliga träningstiden före bytet kan vi inte utsluta att väntvärdet är oförändrat (på 5% signifikansnivå). Det är möjligt att vi skulle ha förkastat H_0 om n hade varit större.

↓ antal slumpmässigt utvalda kunder efter bytet
 För att utföra testet antar vi att träningsstiderna är flyttat normalfördelade och att observationerna är oberoende av varandra (tex att man inte räknar med kompisar som tränas tillsammans och därmed gymmar lika länge).

b) Jämföra två grupper - parade data (före och efter bytet) samma persons
 $H_0: \mu_2 - \mu_1 = 0$
 $H_A: \mu_2 - \mu_1 \neq 0$
 Lat H_1 vara träningsstid före bytet och H_2 väntvärdet för vara väntvärdet för träningsstid efter bytet.

$T = \frac{D - d_0}{s_D/\sqrt{n}}$

Person	Tid innan bytet X_1	Tid efter bytet X_2	Tid efter - Tid innan $D = X_2 - X_1$	$D - d_0$	$(D - d_0)^2$
1	60	72	72-60 = 12	12-19,4 = -7,4	(-7,4) ² = 54,76
2	39	49	49-39 = 10	10-19,4 = -9,4	(-9,4) ² = 88,36
3	46	78	78-46 = 32	32-19,4 = 12,6	12,6 ² = 158,76
4	59	95	95-59 = 36	36-19,4 = 16,6	16,6 ² = 275,56
5	33	40	40-33 = 7	7-19,4 = -12,4	(-12,4) ² = 153,76

$D = \frac{\sum_{i=1}^n D_i}{n} = \frac{19 + 10 + 32 + 36 + 7}{5} = \frac{104}{5} = 20,8$
 $s_D = 19,4$ minuter



Datum: (Date YYYY-MM-DD)	2023-03-20	Kurs/Kurskod: (Course/Course code)	ST1101	Sidnr.: (Page no.)	5
Anonymiseringskod (Anonymization code)	3 1 1 - 0 0 0 4 - L C B				

fortsättning
4.b)

Uppg.nr.:
(Task no.)

4b), 5a)

$$\bar{D} = 19,4 \text{ minuter}$$

$$s_D^2 = \frac{54,76 + 88,36 + 158,76 + 275,56 + 153,76}{5-1} = 182,8$$

$$s_D = \sqrt{182,8} = 13,5204$$

$$T = \frac{\bar{D} - d_0}{s_D / \sqrt{n}}$$

$$T = \frac{19,4 - 0}{13,5204 / \sqrt{5}} = \frac{19,4}{6,0465} = 3,208475$$

Lärarens kommentar:
(Teacher's note)

Hypotestest på 5% signifikansnivå

$$t_{crit} = t_{0,05/2, 5-1} = 2,776$$

4b.) 12

Förkastar H_0 om $|T| \geq t_{crit}$

$$|T| = 3,20847 > 2,776 = t_{crit} \quad \text{Förkastar } H_0$$

Vi förkastar H_0 på 5% signifikansnivå, resultatet tyder på att bytet av spellistor har påverkat längden på passen för gymmets kurser.

Antaganden: oberoende inom grupperna, mellan personerna, och normalfördelade skillnader (differenser).

5 a) $\widehat{\text{Salary}} = 83,312 + 0,650 \cdot \text{yrs. since. phd}$
 $\widehat{\text{Salary}} = 83,312 + 0,650 \cdot 23 = 98,262 \quad (\cdot 1000 \text{ kr})$

$$\hat{y}_* \pm t_{0,05/2, 397-2} \cdot \sqrt{\frac{s_e^2}{n} + s_{b_1}^2 \cdot (x_* - \bar{x})^2 + s_e^2}$$

$$98,262 \pm 1,966 \cdot \sqrt{\frac{23,8^2}{397} + 0,092^2 \cdot (23 - 22,31)^2 + 23,8^2}$$

395 frihetsgrader finns inte i E-fördelningens tabell. Jag utläser från 400 frihetsgrader. $395 \approx 400$

$$98,262 \pm 1,966 \cdot \sqrt{567,8708}$$

$$98,262 \pm 1,966 \cdot 23,83$$

$$98,262 \pm 46,84986$$

$$(51,41214, 145,11119)$$

→ månadslön i tusen kr

EH 95% prediktionsintervall för salary för en professor som fick doktorsexamen för 23 år sedan är mellan 51,41 och 145,11 tusen kr. (51412,14 kr, 145111,9 kr).

5a.) 8

Poäng:
(Points)

(394 frihetsgrader finns ej i t-fördelningens tabell, jag kollar på 400 frihetsgrader) $394 \approx 400$

$$T = \frac{592}{616030 - 0} = 1,6916$$

$$T = b_2 - R_2^{(0)}$$

$t_{crit} = t_{0,05/2, 397-3}$
 $t_{crit} = 1,966$

Förkastar H_0 om $|T| \geq t_{crit}$
 $|T| = 1,6916 < 1,966 = t_{crit}$
 Förkastar inte H_0

Vi kan utföra hypotestestet inte förkastar H_0 (att variablen sexhåle inte är signifikant förklarande) på 5% signifikansnivå.
 Det är möjligt att sexhåle inte är en signifikant förklarande variabel.

c) För att beräkna p-värdet för testet använder jag mig av en normalapproximation. För att göra det behövs $(n-k-1) \geq 30$
 $n-k-1 = 397-2-1 = 394 > 30$

$t_{0,025, 394} \rightarrow Z_{0,025} = 1,96$

$p\text{-värde} = 2 \cdot P(Z \geq T) = 2 \cdot P(Z > 1,69) = 2 \cdot (1 - P(Z \leq 1,69))$
 $= 2 \cdot (1 - 0,9545) = 2 \cdot 0,0455$
 $= 0,09$

Sannolikheten att få ett tofs minst så extremt som 1,6916 är $0,09$.

$0,09 > 0,05$ (5% signifikansnivå) Förkastar ej H_0 .
 Även p-värdet för testet tyder på att vi inte kan förkasta att sexhåle inte är en signifikant förklarande variabel.

5c) 5

5d) 2

d) Justerat R^2 är något högre för modellen med sexhåle som en förklarande variabel (jag bortser från R^2 eftersom det värdet alltid blir högre med fler förklarande variabler). Skillnaden i justerat R^2 mellan modellerna är dock ganska liten (0,11 i första modellen och 0,178 i den andra modellen). Dit tyder på att variablen sexhåle inte bidrar till att förklara så mycket mer av variablen i salary. Eftersom även hypotes-testet med både teststatistikan och p-värdet visade att vi inte kan förkasta H_0 , att sexhåle inte är en signifikant förklarande variabel, på 5% signifikansnivå så verkar inte variablen sexhåle tillföra särskilt mycket till modellen. Den enkla linjära regressionen utan sexhåle kan därför vara att fördra.



Rules in the examination hall

- Follow the invigilator's instructions.
 - Bags and outerwear must be placed at the designated place.
 - Place your ID document clearly visible on the table in front of you.
 - No student may leave the examination hall for the first 30 minutes.
 - Only one student at a time may visit the toilet. Before visiting the toilet, write your name and time on the intended list. After the toilet visit, enter the time on the list again.
 - Electronic equipment such as a mobile phone or Smartwatch must be switched off and placed at the designated place.
 - During the exam, silence applies – you are not allowed to talk, or otherwise communicate, with other students during the exam.
 - Before submitting the examination documents, remember to write the page number, anonymization code, and date on all papers.
- Please do not hesitate to ask the invigilator if anything is unclear. Good luck!

Regler i skrivsalen

- Följ tentamensvårds anvisningar.
 - Väskor och ytterkläder ska placeras på anvisad plats.
 - Placera ID-handling väl synlig på bordet framför dig.
 - Ingen student får lämna skrivsalen under de första 30 minuterna.
 - Endast en student i taget får besöka toaletten. Vid toalettbesök skriv ditt namn och klockslag på avsedd lista. Efter toalettbesök ska du åter ange klockslag på listan.
 - Elektronisk utrustning som mobiltelefon eller Smartwatch ska vara avstängd och placerad på anvisad plats.
 - Under tentamen gäller tystnad – det är förbjudet att prata, eller på annat sätt kommunicera, med andra studenter under pågående tentamen.
 - Innan tentamenshandlingarna lämnas in; skriv sidnummer, anonymiseringskod och datum på alla inlämnade papper.
- Om något är oklart – fråga gärna tentamensvården. Lycka till!!!