



Försättsblad till skriftlig tentamen vid Linköpings Universitet

Datum för tentamen	2011-01-21
Sal (1) Om tentan går i flera salar ska du bifoga ett försättsblad till varje sal och <u>ringa in</u> vilken sal som avses	TER3
Tid	8-12
Kurskod	732G81
Provkod	TENT
Kursnamn/benämning	Statistik
Provnamn/benämning	Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	9
Jour/Kursansvarig Ange vem som besöker salen	Karl Wahlin
Telefon under skrivtiden	0709-719096
Besöker salen ca kl.	10
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Carita Lilja, 1463, carita.lilja@liu.se
Tillåtna hjälpmedel	Valfri räknedosa, formelsamling fri från anteckningar
Övrigt	
Vilken typ av papper ska användas, rutigt eller linjerat	
Antal exemplar i påsen	

Tentamen

Linköpings Universitet, Institutionen för datavetenskap, Statistik

Kurskod och namn:	732G81 Statistik för internationella ekonomer
Datum och tid:	2011-01-21, 8-12
Jourhavande lärare:	Kalle Wahlin
Tillåtna hjälpmedel:	Valfri räknedosa, formelsamling fri från anteckningar.
Betygsgränser:	Tentamen omfattar totalt 40p. Godkänt från 24p, väl godkänt från 32p. Siffrorna i uppgifterna är påhittade.

Redovisa och motivera tydligt alla dina lösningar!

Uppgift 1 (9p)

Diametern på en slumpmässigt vald pizza från en viss pizzeria, A , kan betraktas som en normalfördelad slumpvariabel med väntevärde μ och standardavvikelse σ . 6 pizzor kontrollmäts och följande diametrar (i cm) erhålles:

31 32 30 31 29 30

- Bestäm medelvärde och standardavvikelse för genomsnittsdiametern på pizzor från den studerade pizzerian. (1p)
- Bestäm ett 95% konfidensintervall för genomsnittsdiametern på pizzor från den studerade pizzerian. (2p)
- Är diametern på pizzorna från den studerade pizzerian i genomsnitt större än 30 cm? Ställ upp hypoteser, genomför hypotesprövningen och dra slutsatser med ord. Använd 5% signifikansnivå. (2p)
- Vid en konkurrerande pizzeria, B , drar man ett slumpmässigt urval om 7 pizzor. I detta stickprov bestämmer man genomsnittsdiametern till 31.5 cm och standardavvikelsen till 2.11 cm. Har pizzor från pizzeria B i genomsnitt samma diameter som pizzor från pizzeria A eller ej? Besvara frågan genom att bilda ett lämpligt dubbelsidigt konfidensintervall. Använd 95% konfidensnivå och dra slutsatser med ord. (3p)
- Ange vilka förutsättningar konfidensintervallen och hypotesprövningen baseras på. (1p)

Uppgift 2 (4p)

Man vill undersöka andelen röstberättigade svenskar som vill avveckla kärnkraften. I ett OSU om 1200 personer uppges 612 att de vill se kärnkraften avvecklad före år 2030.

- a) Pröva hypotesen att majoriteten av alla röstberättigade svenskar önskar avveckla kärnkraften mot att så inte är fallet. Svara genom att ange hypotesprövningens p-värde och dra slutsats på 1% signifikansnivå. (2p)
- b) Antag att man vill använda detta OSU som en förberedande studie för en större undersökning. Om man i denna större undersökning i ett 99% konfidensintervall maximalt accepterar en felmarginal om 0.01, hur många personer skulle man då behöva ta med i urvalet? (2p)

Uppgift 3 (3p)

En importör av frukostflingor i 500-gramspaket kontrollerar de leveranser om 10000 paket man mottar genom att kontrollväga ett slumpmässigt urval om 50 paket ur varje parti. Vid en viss leverans uppmäts genomsnittsvikten till 498.8 gram bland 50 slumpmässigt utvalda paket.

Beräkna approximativt sannolikheten att få en genomsnittsvikt bland 50 slumpmässigt valda paket som är så låg som 498.8 gram (eller lägre) om det är sant att genomsnittsvikten per paket i hela partiet är 500 gram och standardavvikelsen mellan paketvikter är 4.0 gram (vilket leverantören hävdar).

Uppgift 4 (5p)

Vinstplanen för en normalserie om 16 miljoner Trisslotter ser ut på följande sätt.

Antal	Vinst	Antal	Vinst
8	2500000	1200	750
8	1000000	1600	500
40	250000	4000	250
8	200000	3600	200
16	100000	10000	150
16	20000	75200	100
320	10000	238400	75
1120	2000	1672800	50
1680	1000	1336000	25

Resten av lotterna i serien är nitlotter. Trisslotter kostar 25 kronor styck.

- a) En person har vunnit 25 kronor på en Trisslott och erbjuds att antingen få en ny lott eller de 25 kronorna. Vilket bör personen välja, om hon bortser från spänningsmomentet? Visa uträkning och förklara. (2p)
- b) En annan person köper 100 Trisslotter. Vad är sannolikheten för att minst 25 av dessa är vinstlotter (25 kronor eller mer)? (3p)

Uppgift 5 (4p)

Kretskort som tillverkas i en fabrik kan ha två typer av fel, A och B. Sannolikheten för att produkten ska ha fel A är 30% och sannolikheten för fel B är 40%. De två typerna av fel är oberoende.

Vi drar slumpmässigt ett kretskort ur företagets produktion.

- a) Vad är sannolikheten för att kretskortet har båda typerna av fel? (2p)
- b) Vad är sannolikheten för att kretskortet inte har något av felen? (2p)

Uppgift 6 (4p)

Man har gjort två urval, det ena bestående av 400 bland 8000 TekFak-studenter och det andra bestående av 600 bland 12000 FilFak-studenter, och tillfrågat de utvalda studenterna om genomsnittlig litteraturkostnad (i hundratals kronor) under den gångna terminen. Följande resultat erhöles.

	TekFak	FilFak
Medelvärde	4.84	3.27
Standardavvikelse	2.94	1.12

Beräkna ett 95% konfidensintervall för den genomsnittliga litteraturkostnaden per studerande bland de 20000 studenterna.

Uppgift 7 (3p)

Bärgningsfirman i en region noterar för ett antal incidenter under vinterhalvåret om de förolyckade bilarna haft dubbdäck, friktionsdäck eller sommardäck. För varje incident har också väglaget noterats, enligt A: isgata, B: snömodd eller C: gott skick.

	A	B	C
Dubbdäck	38	30	34
Friktionsdäck	31	26	28
Sommardäck	31	10	8

Finns det något samband mellan typ av däck på det förolyckade fordonet och väglag? Besvara frågan genom en lämplig hypotesprövning på 5% signifikansnivå.

Uppgift 8 (2p)

Ta för vart och ett av följande påståenden ställning till om yttrandet är korrekt eller ej. Kort motivering ska ges.

- a) Allmänt gäller i 95% konfidensintervall för en populationsandel, att felmarginalen är 0.05. (0.5p)
- b) Medianen i ett datamaterial ligger alltid mitt emellan den undre och den övre kvartilen. (0.5p)
- c) Standardavvikelsen bland ett antal observationer är alltid mindre än variansen i samma material. (0.5p)
- d) I ett datamaterial som fördelar sig symmetriskt kring sitt medelvärde är medianen = medelvärdet. (0.5p)

Uppgift 9 (6p)

En privatekonom undersöker månatlig inkomst före skatt och månatligt sparande bland 15 slumpmässigt utvalda personer. Följande resultat erhålles.

Inkomst (tkr)	26,5	19,5	24,0	23,0	29,0	18,5	16,0	26,0	41,5	27,0	24,5	14,0	22,0	12,5	8,5
Sparande (tkr)	0,5	0,0	1,2	0,8	0,5	2,0	0,3	0,2	5,0	2,0	1,5	0,2	0,4	0,8	0,1

- Beräkna korrelationskoefficienten mellan månadsinkomst och månadssparande. (1.5p)
- Beräkna a och b i en enkel linjär regressionsmodell. (3p)

Nedan följer en delvis censurerad Minitabanalys av sambandet mellan månadsinkomst och månadssparande.

Regression Analysis: Sparande (tkr) versus Inkomst (tkr)

The regression equation is

Predictor	Coef	SE Coef	T	P
Constant		0,6962		
Inkomst (tkr)		0,03002		

S = 0,929253 R-Sq = 50,8% R-Sq(adj) = 47,0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	11,568	11,568	13,40	0,003
Residual Error	13	11,226	0,864		
Total	14	22,793			

Unusual Observations

Obs	Inkomst (tkr)	Sparande (tkr)	Fit	SE Fit	Residual	St Resid
9	41,5	5,000	3,202	0,639	1,798	2,67RX

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.

Predicted Values for New Observations

New Obs	Fit	SE Fit	95% CI	95% PI
1	1,389	0,259	(0,829; 1,948)	(-0,695; 3,473)

Values of Predictors for New Observations

New Obs	Inkomst (tkr)
1	25,0

- c) Vilket av följande påståenden stämmer bäst? Svara bara med nummer. Ingen motivering eller uträkning behöver lämnas. (1.5p)
- I. 47% av variationen i månadssparande förklaras av månadslön.
 - II. Regressionsanalys är i detta fall olämplig analysmetod eftersom variablerna är kvalitativa.
 - III. Med 95% säkerhet ligger det månatliga sparandet för en viss person med månadsinkomsten 25000 kronor mellan 829 och 1948 kronor.
 - IV. $t = 3.66$ varför det föreligger ett statistiskt säkerställt samband mellan månadsinkomst och månadssparande på 5% nivå.
 - V. Intercepttermen har i detta fall en logisk tolkning eftersom $y = 0$ ingår i observationsområdet.
 - VI. Regressionsanalysen är ogiltig eftersom Minitab varnar för en outlier.