

Tentamen

Linköpings universitet, Institutionen för datavetenskap, Statistik

Kurskod och namn:	732G81 - Statistik för internationella ekonomer
Datum och tid:	1 februari 2016, 8-12
Jourhavande lärare:	Isak Hietala
Tillåtna hjälpmedel:	Valfri räknedosa, formel- och tabellsamling utan anteckningar.
Betygsgränser:	Tentamen omfattar totalt 40p. G från 24p och VG från 32p. Siffrorna i uppgifterna är påhittade.

Redovisa, tolka och motivera tydligt alla dina lösningar!

Uppgift 1 (5p)

Varje år på kvarteret Labradoren ordnas en tävling i blomsterodling, där första pris går till den boende som har den största medelhöjden på sina solrosor i slutet av sommaren. Bland alla tävlandes rosor väljs 10 stycken ut slumpmässigt som den boendes bidrag till tävlingen. Följande höjd mättes på de 10 utvalda blommorna bland de 35 som Herr Eriksson hade i sin odling.

82 93 59 58 69 89 52 63 85 76

1. Beräkna medelhöjden av blommorna. (1p)
2. Beräkna standardavvikelsen av blommornas höjd. (1p)
3. Herr Eriksson påstår sig att om man staplar alla hans rosor på varandra, kommer de sträcka sig 20 meter upp i luften. Beräkna ett nedåt begränsat konfidensintervall för den totala höjden av rosorna och kontrollera Herr Erikssons påstående. (3p)

Uppgift 2 (5p)

I genomsnitt köper ungefär 89.2 procent av alla ekonomistuderer en Texas Instruments miniräknare. Vad är sannolikheten att av...

1. ... 5 studenter, minst 2 har en Texas-miniräknare? (2p)
2. ... 1000 studenter, fler än 500 har en Texas-miniräknare? (3p)

Uppgift 3 (10p)

I en klass undersöktes sambandet mellan ålder och tentamensresultat och följande regressionsutskrift erhöles.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,604 ^a	,365	,351	2,6962

a. Predictors: (Constant), Ålder

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	187,817	1	187,817	25,837	,000 ^b
	Residual	327,119	45	7,269		
	Total	514,936	46			

a. Dependent Variable: Resultat

b. Predictors: (Constant), Ålder

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3,752	1,911		1,964	,056
	Ålder	,422	,083	,604		,000

a. Dependent Variable: Resultat

1. Skriv upp den skattade stickprovsmodellen. (2p)
2. Tolka den skattade lutningsparametern b_1 i ord. (1p)
3. Testa på 5 procents signifikans huruvida det finns ett statistiskt säkerställt samband mellan ålder och resultat. (2p)
4. Identifiera och tolka förklaringsgraden. (1p)
5. Beräkna en punktskattning för tentamensresultatet av en student som är 25 år gammal. (1p)
6. Beräkna ett 95-procent konfidensintervall för studenter som är 25 år gamla. (3p)

Uppgift 4 (7p)

I en marknadsundersökning för en ny produkt visas tre olika sorters förpackningar, A, B och C. De svarande blev indelade i tre åldersgrupper och fick besvara frågan vilken förpackning de tyckte bäst om. Följande resultat erhöles.

	A	B	C
Yngre	10	20	38
Medel	12	20	39
Äldre	19	4	3

1. Företaget är intresserad att undersöka ifall de olika åldersgrupperna har annorlunda preferenser om förpackning. Vilket test är lämpligast att utföra? Motivera. (1p)
2. Undersök nu huruvida det finns skillnader i förpackningspreferens mellan de olika grupperna med en lämplig hypotesprövning. Använd 5 procents signifikansnivå. (2p)
3. Undersök nu huruvida det finns skillnader i **andelen** yngre och medelålder personer som föredrar förpackning A. Besvara frågeställningen genom en lämplig hypotesprövning och använd 1 procents signifikansnivå. (3p)
4. Vilka antaganden krävs för att kunna utföra testet i fråga 3? (1p)

Uppgift 5 (8p)

Mätningar av ammoniakutsläpp (g/m^3) görs varje år på slumpmässigt valda platser i övre och nedre Norrland uppdelat på tre områden; från stall, betesmarker och gruvor.

Område	Övre (f)	Övre (medelutsläpp)	Nedre (f)	Nedre (medelutsläpp)
Stall	75	458.12	94	782.69
Betesmark	98	195.76	92	254.81
Gruvor	32	152.54	12	98.65

1. Beräkna medelutsläppet av ammoniak i de båda regionerna. Vilken region verkar ha högst utsläppsnivå? (2p)
2. Beräkna ett dubbelsidigt konfidensintervall på 5 procents signifikansnivå över skillnaden i medelutsläpp mellan de två regionerna. (3p)
3. Beräkna medelutsläppet genom att kompensera för att områdesfördelningen är olika i de båda regionerna. Jämför med de medeltal som beräknades i 1. (3p)

Uppgift 6 (5p)

I en utvärdering av en statistikkurs gavs följande helhetsbetyg på kursen.

Betyg	Antal
1	0
2	5
3	19
4	34
5	24

1. Visualisera tabellen med ett lämpligt diagram. (2p)
2. Beräkna medianen samt första och tredje kvartilen i datamaterialet. (2p)
3. Bland sex slumpmässigt utvalda studenter som läste kursen, vad är sannolikheten att alla sex gav kursen betyget 5? (1p)