



Försättsblad Prov Original

Kurskod	Provkod	Tentamensdatum
S T 0 0 1 G	T 1 0 2	2 0 1 8 - 1 1 - 0 7
Kursnamn	Statistik GR (A), Grundkurs	
Provnamn	Skriftlig tentamen	
Ort	Östersund	
Termin	H18	
Ämne	Statistik	

Mittuniversitetet
EJV, Avdelningen för ekonomivetenskaper och juridik

Tentamen statistik A1 15 hp

Dag: 2018-11-07

Tid till förfogande: 5 timmar

OBLIGATORISK LÄSNING

Hjälpmedel: Miniräknare (delas ut av tentamensvakterna)
Tabell- och formelsamling

Anvisningar: Redovisa tydligt tankegången i lösningarna. Visa **samtliga** uträkningar där inget annat är angivet.

Redovisa varje uppgift på separat ark.

Betygsgränser

A 45-50

B 40-44

C 35-39

D 30-34

E 25-29

F 0-24

Fx ges inte i kursen. Du kan alltså **inte** få komplettering upp till E. Detta är ett principbeslut inom ämnet företagsekonomi där statistik kursen ingår.

Uppgift 1**(3 p)**

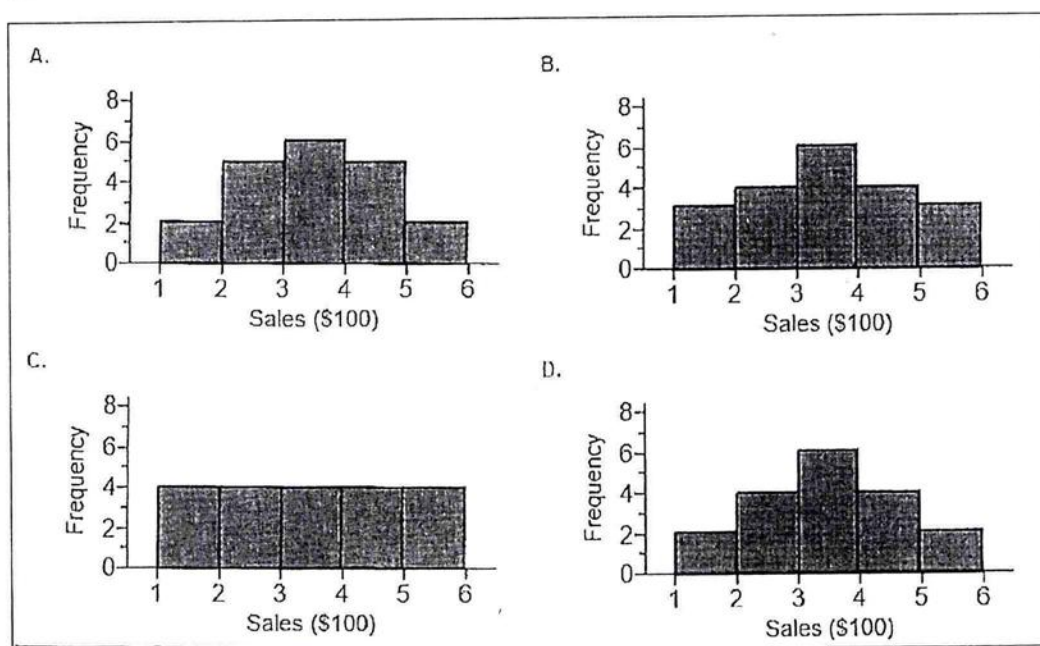
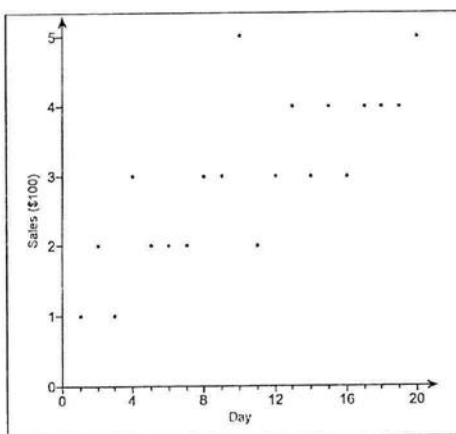
Spridningsdiagrammet visar daglig försäljning i tusentals US-dollar för ett nyöppnat fik.

Poängfördelning: rätt svar = 1 p, fel svar = -1 p och obesvarad fråga = 0 p.

(Lägsta poäng på frågan är 0 p. Du ska inte motivera ditt svar eller visa uträkningar)

SVARA PÅ LÖSNINGSBLAD

- A. Studera spridningsdiagrammet: Vilket av histogrammen (A-D) passar bäst för att beskriva försäljningen?



- B. Vad syns tydligt i spridningsdiagrammet men inte i histogrammet? Välj ett av alternativen a-d!
- a. Att försäljning har en bimodal och symmetrisk fördelning.
 - b. Att försäljning har en unimodal och symmetrisk fördelning.
 - c. Att sambandet mellan försäljning och dag är negativt.
 - d. Att sambandet mellan försäljning och dag är positivt.
- C. Vad syns tydligare i histogrammet men inte i spridningsdiagrammet? Välj ett av alternativen a-d!
- a. Att försäljning har en unimodal och symmetrisk fördelning.
 - b. Att sambandet mellan försäljning och dag är positivt.
 - c. Att sambandet mellan försäljning och dag är negativt
 - d. Att försäljning har en unimodal och asymmetrisk fördelning.

UPPGIFT 1: SVARA PÅ LÖSNINGSBLAD

Uppgift 2

(3 p)

Redogör för urvalsmetoden klusterurval. Ge även exempel.

Uppgift 3

(15 p)

I en finsk studie om 150 kvinnor delades de slumpmässigt in i en av de tre behandlingarna. Samtliga drycken såg identiska ut och kvinnorna visste inte vilken dryck de fick.

Frågeställningen: Minskar urinvägsinfektioner med hjälp av tranbärsjuice bland kvinnor?

Resultatet av studien presenteras i korstabellen nedan.

Beräkningarna ska redovisas och svaren tydligt framgå.

	Tranbärsdryck	Lactobacillus dryck	Placebo	Totalt
Minst en infektion inom sex månader	8	20	18	46
Ingen infektion inom sex månader	42	30	32	104
Totalt	50	50	50	150

- A. Studera korstabellen ovan och ange samtliga mätvärden för vart och ett av variablerna.
- B. Beräkna statistiska effekten av x på y i form av *differens i andel*. Tänk på studiens frågeställning. (Du ska inte beräkna konfidensintervall för din jämförelse)
- i) Ange vilken jämförelse du vill göra.
ii) Ange hur stor skillnaden blir i din jämförelse.
- C. Avgör med hjälp av ett chitvå-test om urinvägsinfektioner minskar med hjälp av tranbärsjuice bland kvinnor i populationen. Hypotesformuleringar, diskussion av förutsättningar samt relevanta beräkningar och konklusion ska redovisas.
- D. Vad kallas denna studiedesign?

Uppgift 4

(6 p)

Ett företag som säljer livförsäkringar vill skatta premiens belopp med hjälp av inkomst. Båda variablerna är mätta i 1000-tals USD. $n = 20$ hushåll

Uträkningar ska redovisas och svaren tydligt framgå.

Resultatet av analysen redovisas nedan:

Coefficients					
Model		Unstandardized Coefficients		Standardized	Sig.
		B	Std. Error	Beta	
1	(Constant)	-,862	1,923		,659
	Inkomst	,254	,007	,993	,000

- Är sambandet mellan premiens belopp och inkomst positivt eller negativt? Motivera svaret utifrån SPSS-utskriften!
- Ange regressionsekvationen!
- Tolka värdet av b_1 i just detta sammanhang?
- Beräkna residual för en familj där inkomsten är 90 000 USD och premiens belopp är 25 000 USD.
- Tolka värdet på residualen från D-uppgiften.

Uppgift 5

(5 p)

I Nya Zeeland har 3,24 % av européerna och 1,77 % av maorierna blodtyp AB. En blodbank i ett distrikt där befolkningen till 85 % är europeisk och till 15 % maorisk, behöver veta hur mycket blod av typ AB man behöver ha i lager. Några frågor infinner sig:

Uträkningar ska redovisas och svaren tydligt framgå.

- A. Vilken typ av sannolikhet är talet 3,24 %?
- B. Hur många procent av befolkningen i distriktet har blod av typ AB?
- C. Vilken typ av sannolikhet efterfrågas i fråga B?

Uppgift 6

(7 p)

Vid köp av ny bil får kvinnor i genomsnitt 15 000 kr i rabatt från listpriset med standardavvikelse 3 500 kr och motsvarande rabatt för män är i genomsnitt 10 000 kr med standardavvikelse 3 000 kr. Den erhållna rabatten är en slumpvariabel (approximativt normalfördelad) och oberoende mellan könen.

Uträkningar ska redovisas och svaren tydligt framgå.

OBS! För full poäng krävs dessutom att du definierar dina variabler samt anger vad du beräknar med ett skrivsätt typiskt för uppgiften.

Hur stor är sannolikheten att en slumpvis vald man får mer rabatt än en slumpvis vald kvinna?

Uppgift 7

(11 p)

En tärning kastas 210 gånger med följande resultat, medelvärde 3,638 och standardavvikelse 1,692.

Följande fråga ska besvaras med hjälp av sannolikhetsbaserade analyser.
”Kan man påstå att tärningen är symmetrisk?”

För delfrågorna A, B och C

Välj ett av alternativen a-e.

Poängfördelning: 1 rätt=1p, 2 rätt=2p och 3rätt=4p.

(Inga minuspoäng delas ut vid fel svar och du ska heller inte motivera ditt svar eller visa uträkningar)

A) Vilket par av noll- och mothypotes ska använda?

- | | | |
|----------------------------|-----|----------------------------|
| a) $H_0 : \mu = 3,638$ | och | $H_1 : \mu \neq 3,638$ |
| b) $H_0 : p = 0,500$ | och | $H_1 : p \neq 0,500$ |
| c) $H_0 : \bar{x} = 3,638$ | och | $H_1 : \bar{x} \neq 3,638$ |
| d) $H_0 : \beta = 3,500$ | och | $H_1 : \beta \neq 3,500$ |
| e) $H_0 : \mu = 3,500$ | och | $H_1 : \mu \neq 3,500$ |

B) Vad är testvariabelns observerade värde, dvs test statistics?

- a) ca 17,250
- b) ca 12,150
- c) ca 2,150
- d) ca 1,182
- e) ca 0,082

C) Vad är p-värdet för det observerade värdet i denna hypotesprövning?

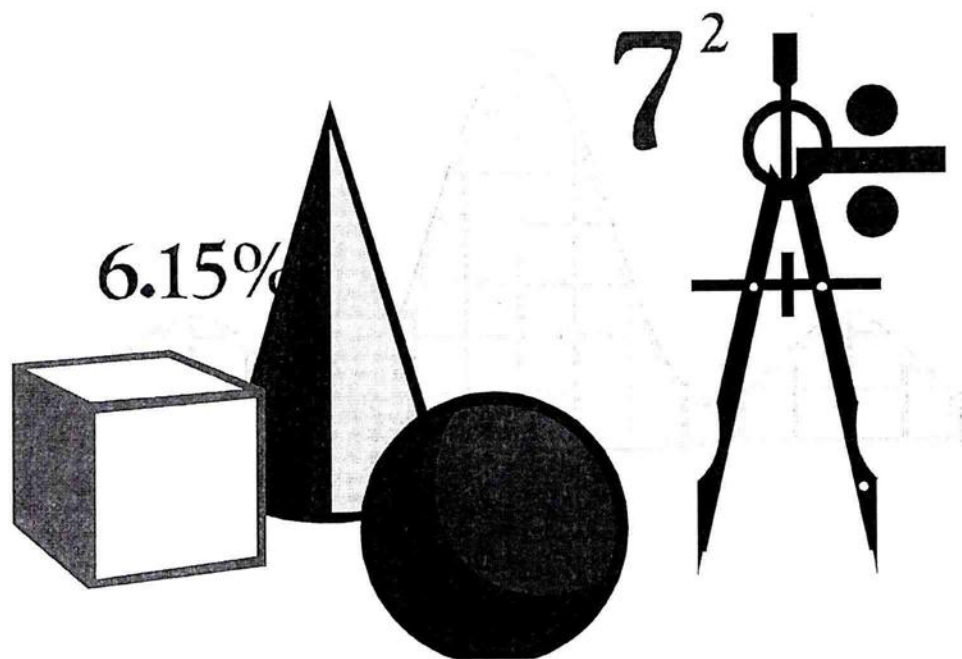
- a) mer än 0,20
- b) mellan 0,10 – 0,20
- c) mellan 0,05 - 0,09
- d) mellan 0,01 – 0,02
- e) mindre än 0,01

D) Beräkna konfidensintervall, 95 %, för parametern. **Uträkningar och använda formler ska redovisas samt svaren tydligt framgå.**

E) Besvara nu frågan med hjälp av resultaten av de sannolikhetsbaserade analyserna (test och konfidensintervall): Kan man påstå att tärningen är symmetrisk? **Svara ja eller nej och motivera ditt svar!**

F) Om du vill halvera din felmarginal (margin of error) i konfidensintervallet i D-uppgiften, hur stort n behövs? **Motivera svaret!**

Lämna endast in lösningar



Formelsamling i statistik, A1, 15 hp

2016-01-25

Innehåll

BETECKNINGAR	2
KAP 1 DATA AND DECISIONS	2
KAP 3 DISPLAYING AND DESCRIBING QUANTITATIVE DATA	2
KAP 4 CORRELATION AND LINEAR REGRESSION.....	3
KAP 5 RANDOMNESS AND PROBABILITY	3
KAP 6 RANDOM VARIABLES AND PROBABILITY MODELS.....	3
KAP 7 THE NORMAL DISTRIBUTION.....	4
KAP 9+11 SAMPLING DISTRIBUTIONS.....	4
KAP 9-15 INFERENCE.....	5

Beteckningar

Stora bokstäver, X, Y etc., betecknar slumpvariabler.

Små bokstäver, x, y etc., betecknar faktiska värden på observationer.

Populationsstorleken betecknas med N .

Stickprovsstorleken betecknas med n .

Grekiska bokstäver betecknar populationens parameter. Ex σ , som betecknar populationens standardavvikelse

Latinska bokstäver betecknar skattningen av parametern. Den skattas utifrån stickprovet. Ex s , som betecknar stickprovets standardavvikelse.

	<u>Parameter (i populationen/ sannolikhetsfördelningen)</u>	<u>Parameterskattning (statistika) (i stickprovet)</u>
Medelvärde	μ	\bar{y} eller \bar{x}
Varians	σ^2	s^2
Standardavvikelse	σ	s
Andel, proportion	p	\hat{p}
Generellt	θ	$\hat{\theta}$

Kap 1 Data and Decisions

Five W' + How: Who, What, When, Where, Why and How

Kap 3 Displaying and Describing Quantitative Data

$$\text{Range} = \text{Max} - \text{Min}$$

$$\text{IQR} = Q3 - Q1$$

$$\text{Outlier Rule-of-Thumb: } y < Q1 - 1.5 \times \text{IQR} \text{ or } y > Q3 + 1.5 \times \text{IQR}$$

$$\bar{y} = \frac{\sum y}{n}$$

$$s = \sqrt{\frac{\sum (y - \bar{y})^2}{n - 1}}$$

$$z = \frac{y - \mu}{\sigma} \text{ (model based)}$$

$$z = \frac{y - \bar{y}}{s} \text{ (data based)}$$

