



## Försättsblad Prov Original

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

Mittuniversitetet  
EJV, Avdelningen för ekonomivetenskaper och juridik

Tentamen statistik A1 15 hp

Dag: 2018-12-12  
Tid till förfogande: 5 timmar

---

## OBLIGATORISK LÄSNING

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

**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 statistikkursen ingår.

**Uppgift 1****(6 p)**

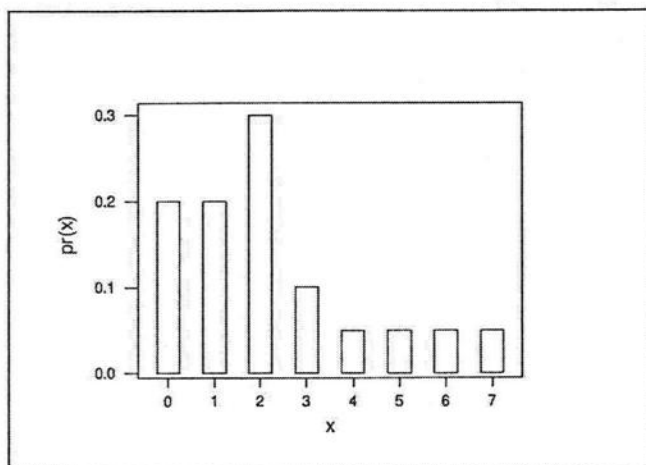
Denna uppgift handlar om lärandemålet sannolikheter och sannolikhetsfördelningar.

Uträkningar ska redovisas och svaren tydligt framgå.

- A. Slumpvariabeln  $X$  är normalfördelad med  $\mu = 10$  och  $\sigma = 8$ . Beräkna  $P(-4,0 < x < 18,0)$ !
- B. Slumpvariabeln  $X$  är normalfördelad med  $\mu = -15$  och  $\sigma = 10$ .  
Ersätt frågetecknet i uttrycket  $P(x < ?) = 0,90$

**Uppgift 2****(6 p)**

Slumpvariabeln  $X$  är modell för antalet förhandsanmälda till en viss båtut. Den som ansvarar för båtturen har angivit en subjektiv sannolikhetsfördelning för hur många som kommer att anmäla sig en viss dag. Sannolikhetsfördelningen definieras enligt diagrammet.



Vi ser t ex. att sannolikheten för att vi får värdet 1 är 0,2 om vi gör en slumpvis observation av  $X$ .

Uträkningar ska redovisas och svaren tydligt framgå.

- A. Vad är sannolikheten för att det blir fler än två anmälda till denna tur?
- B. Vad är sannolikheten för att antalet anmälningar är fler än två, givet att det åtminstone är någon som anmält sig.

Anta att sannolikhetsfördelningen är oförändrad under sex dagar. Den som ansvarar för båtturen definierar en ny slumpvariabel  $Y$ . Variabeln  $Y$  är antalet dagar av dessa sex som det blir fler än två anmälningar.

C. Vilka möjliga variabelvärden har denna slumpvariabel  $Y$ ?

### Uppgift 3

(6 p)

Den här frågan handlar om typ av variabel, variabelvärde (mätvärde) samt datanivå (skalnivå).  
Tänk på att mätvärden ska vara rimliga.

- A. Hitta på en (och endast en) kvalitativ variabel som handlar om motion eller träning. Du får inte välja variabeln kön. Ange:
- tre tänkbara mätvärden för variabeln
  - datanivå för variabeln
  - enhet om det finns
- B. Hitta på en (och endast en) kvantitativ diskret variabel som handlar om motion eller träning. Du får inte välja variabeln kön. Ange:
- tre tänkbara mätvärden för variabeln
  - datanivå för variabeln
  - enhet om det finns.
- C. Hitta på en (och endast en) kvantitativ kontinuerlig variabel som handlar om motion eller träning. Du får inte välja variabeln kön. Ange:
- tre tänkbara mätvärden för variabeln
  - datanivå för variabeln
  - enhet om det finns.

### Uppgift 4

(2 p)

Ett slumpmässigt stickprov om fem elever från en stor klass och deras provresultat presenteras nedan. Vi ser att sista barnet hade 10 poäng på provet.

10  
13  
16  
11  
10

Uträkningar ska redovisas och svaren tydligt framgå.

Beräkna **variansen** för variabeln antal poäng!

### Uppgift 5

(3 p)

Redogör för självval samt diskutera om urvalsmetoden är bra eller inte?

### Uppgift 6

(8 p)

Här kommer några korta frågor från lärandemålet inferens.

Uträkningar ska redovisas och svaren tydligt framgå.

- A. En stickprovsundersökning om 40 objekt visar att medelvärdet är 60 och standardavvikelsen 10. Beräkna felmarginalen!
- B. En stickprovsundersökning om 40 personer visar att 22 personer tycker bäst om sommaren av alla årstider. Beräkna standard error för stickprovsandel!
- C. En stickprovsundersökning om 300 personer visar att 117 av 150 kvinnor och 90 av 150 män åkte till köpcentra för att fynda på Black Friday. Beräkna felmarginalen för skillnad i andel som åkte till köpcentra för fynda på Black Friday mellan könen.
- D. Vid ett goodness of fit test är  $\chi^2 = 2,783$  och  $df=4$ . Vad blir p-värdet?
- E. Om  $H_A: p < 0,30$ , stickprovsandel = 0,28 och  $n=100$ , vad blir testvariabelns observerade värde, d v s test statistic?
- F. Från B-uppgiften: En stickprovsundersökning om 40 personer visar att 22 personer tycker bäst om sommaren av alla årstider. Vid beräkning av konfidensintervall för p använder vi oftast normalfördelning. Kan vi göra det i denna situation? Motivera svaret!



### Uppgift 7

(7 p)

För mer än 100 år sedan lämnade RMS Titanic hamnen i South Hampton med ca 2 200 personer ombord men lång ifrån alla ombordvarande kom fram till den tilltänkta destinationen, New York. En forskare samlade in data om bland annat överlevnadsstatus och vilken biljett passagerarna köpt (första, andra eller tredje klass). Dessutom fanns en stor besättning ombord. Nedan presenteras resultatet av variablerna överlevnad och klass i en korstabell.

#### Klass \* Överlevnad Crosstabulation

Count

		Överlevnad		Total
		Omkom	Överlevde	
Klass	Besättning	673	212	885
	Första klass	122	203	325
	Andra klass	167	118	285
	Tredje klass	528	178	706
Total	1490	711	2201	

Uträkningar ska redovisas och svaren tydligt framgå och **du får inte slå ihop kategorier för att lösa uppgiften.**

- Rita ett grupperat stapeldiagram (i kursboken kallas den för side-by-side bar chart eller i SPSS för clustered bar chart) i **procent** så att statistiska effekten av *klass* på *överlevnad* framgår. Rita prydligt och var noga med skalan på den horisontella axeln!
- Tolka en (och endast en) av staplarna, i diagrammet A-uppgiften, i just detta sammanhang. Tolkningen får bestå av max 25 ord. (Mer än 25 ord ger poängavdrag)
- Hur stor är den statistiska effekten av *klass* på *överlevnad*, i form av **differens i andel**? Beräkna en (och endast en) differens.
- Tolka talet av värdet på den statistiska effekten i C-uppgiften i just detta sammanhang.

**Uppgift 8**

(12 p)

Studera regressionsutskriften som analyserar sambandet mellan försäljning i 1000-tals USD och antal säljare. Tio butiker valdes slumpmässigt för en viss butikskedja.

Kan vi påstå att försäljning har ett linjärt samband med antal säljare i populationen?

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,965 <sup>a</sup>	,932	,923	1,47701

a. Predictors: (Constant), Number of sales people

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square
1	Regression	238,947	1	238,947
	Residual	17,453	8	2,182
	Total	256,400	9	

a. Dependent Variable: Sales (in \$1000)

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients	
		B	Std. Error
1	(Constant)	8,101	1,021
	Number of sales people	,913	,087

a. Dependent Variable: Sales (in \$1000)

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.

Uträkningar kommer inte att beaktas)

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

- |                            |     |                            |
|----------------------------|-----|----------------------------|
| a) $H_0 : \beta_1 = 0$     | och | $H_1 : \beta_1 \neq 0$     |
| b) $H_0 : b_1 = 0$         | och | $H_1 : b_1 \neq 0$         |
| c) $H_0 : \bar{x} = 0,913$ | och | $H_1 : \bar{x} \neq 0,913$ |
| d) $H_0 : \beta_1 = 0,913$ | och | $H_1 : \beta_1 \neq 0,913$ |
| e) $H_0 : b_0 = 8,101$     | och | $H_1 : b_0 \neq 8,101$     |

B) Vad är testvariabelns observerade värde, d v s test statistic?

- a) ca 8,873
- b) ca 0,087
- c) ca 10,491
- d) ca 11,736
- e) ca 0,913

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) Vad säger konfidensintervallet, i deluppgift D, i just denna kontext.

F) Besvara nu frågan med hjälp av resultaten av de sannolikhetsbaserade analyserna (test och konfidensintervall):

Kan vi påstå att försäljning har ett linjärt samband med antal säljare i populationen vid  $\alpha = 0,05$ ?

**Svara ja eller nej och motivera ditt svar!**

*Lämna endast in lösningar(inte frågorna)*

































