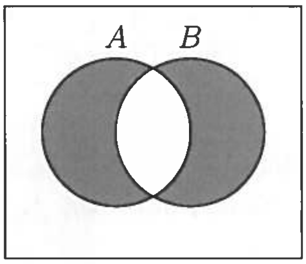
STOCKHOLMS UNIVERSITET
Statistiska institutionen
Jolanta Pielaszkiewicz

TENTAMEN I STATISTIKENS GRUNDER 1
2019-02-19

Skrivtid: 13.00-18.00
Godkända hjälpmedel: Miniräknare, språklexikon.
Tentamen består av fem uppgifter. För full poäng på en uppgift krävs tydliga, utförliga och väl motiverade lösningar.

Uppgift 1. (15 poäng)



a) Är det möjligt att mängden A och B har samma antal individer om gråmärket-området inte är tomt? Om ja, ge exempel på sådana A och B.

b) Beskriva mängden markerad med grå färg med hjälp av A och B. Använd $\cup$, $\cap$, $\setminus$ om behövs.

c) Låt $P(A) = 0.3$ och $P(A \cap B) = 0.1$. Vilket villkor måste B uppfylla så att A och B är oberoende? Motivera svar.

d) Avgör om försök: "en student väljs slumpmässigt ut och personens vikt (avrundad till heltal) noteras" är slumpmässig och, i så fall, om utfallsrummet är ändlig eller oändlig. Vilken skalnivå variabeln befinner sig på?

Uppgift 2. (25 poäng) Hanna vill söka in till Beckmans konstskola och till Konstfack. En konstprofessor har granskat hennes arbetsprover och bedömer sannolikheten att hon kommer att bli antagen till Beckmans till 0.8 och till Konstfack till 0.4.
Sannolikhet att Hanna blir antagen till både Beckmans och Konstfack är 0.3.

a) Vad är sannolikheten för att Hanna **inte** kommer att erbjudas en plats på varken Beckmans eller Konstfack?

b) Professor vet om 5 andra personer som vill söka till Beckmans och Konstfack och har bedömts på samma sätt (fick samma sannolikhet att bli inskrivna). Vad är sannolikheten för att minst 4 av de 6 personerna kommer att **accepteras** till minst en av konstskolorna? Vi antar att personerna är oberoende av varandra.

c) Studenterna på Konstfack har en chans att vinna ett pris för bästa muntliga presentation om några utvalda artister. Man kan få 5000 sek, 2000 sek eller 1000 sek beroende på placering. Det finns 90 elever som deltar i tävlingen. På hur många sätt kan du välja de tre personerna?

d) Är det rimligt att anta normal fördelning för den stokastiska variabeln "antal antagna studenterna"? Motivera svaret.

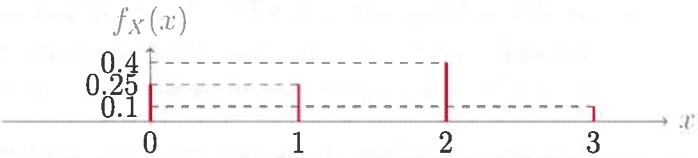
Uppgift 3. (20 poäng) Resultatet från företag A och B antas vara normalfördelad stokastisk variabel. Man bedömer att företag A ger vinst som överstiger 120 000 sek med 0.5% sannolikhet och vinst under 50 000 sek eller förlust med 0.5% sannolikhet. Vi vet också att resultat från företag B har väntevärde 50000 sek och standardavvikelse lika med 100.

a) Vad är det förväntade resultatet för företag A?

b) Vad är förväntade sammanlagda resultatet av de två företag?

c) Kovarians mellan resultaten av företag A och B är 40 000.
   - Beräkna standardavvikelse för resultat för företag A.
   - Beräkna standardavvikelse för det sammanlagda resultatet för företagen.
   - Beräkna korrelation mellan resultaten av företag A och B. Bestäm om korrelation är stark/svag, positiv/negativ.

Uppgift 4. (20 poäng) Sannolikhetsfördelning för efterfrågan på varorna A och B ges av nedanstående tabell och diagram. X och Y betecknar antalet enheter av respektive vara.



y	1	2
$f_Y(y)$	0.75	0.25

a) Bestäm simultana sannolikhetsfördelningen i tabellform för variablerna X och Y så att marginalfördelningen stämmer med informationen ovan och $f_{X,Y}(2,2) = 0.25$.

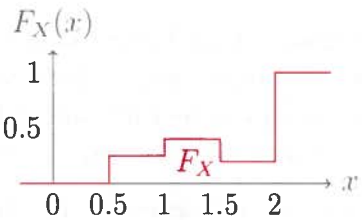
b) Bestäm betingade fördelningarna för efterfråga av A givet att 1 enhet av B efterfrågats.

c) Beräkna förväntade efterfrågan av vara A givet att 1 enhet av B efterfrågats och motsvarande varians.

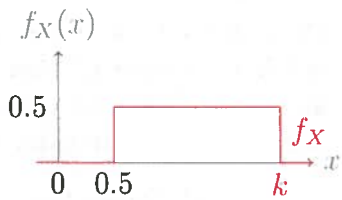
Uppgift 5. (20 poäng)

a) Kan funktionen som presenteras i figur 1 (nedan) vara kumulativ fördelningsfunktion? Motivera svaret.

b) Välj k så att funktionen i figur 2 (nedan) är en täthetsfunktion av kontinuerlig stokastisk variabel. Motivera svaret.



Figur 1: Uppgift 5a)



Figur 2: Uppgift 5b)

c) Beräkna $P(X \geq 7)$ för $X \sim Bin(20, 0.2)$

d) Beräkna $P(X \geq 36)$ för $X \sim Bin(80, 0.4)$

e) Beräkna $P(3 < X \leq 6)$ för $X \sim N(2, \sigma^2 = 9)$

Stockholms universitet

Statistiska institutionen

Rättningsblad

Datum: 19/2-2019

Sal: Värtasalen

Tenta: Statistikens grunder 1

Kurs: Statistikens grunder

ANONYMKOD: 0019-WTN

☒ Jag godkänner att min tenta får läggas ut anonymt på hemsidan som studentsvar.

OBS! SKRIV ÄVEN PÅ BAKSIDAN AV SKRIVBLADEN

Markera besvarade uppgifter med kryss

1	2	3	4	5	6	7	8	9	Antal inl. blad
X	X	X	X	X					14 BC
Lär.ant. 15	25	17	20	13.5					

POÄNG	BETYG	Lärarens sign.
90,5	A	[signature]

SU, STATISTIK

Skrivsal: ________ Anonymkod: 0019-WTN Blad nr: 1

1) a) Ja det är möjligt. Låt Ω vara en hel skolklass. Låt A = "Alla som spelar fotboll"

och låt B = "Alla födda i januari, februari eller mars"

Om 5 personer i klassen spelar fotboll och

5 personer är födda i januari, februari eller mars, så har mängden A och B sammma

/3 - antal individer. och det finns någon som är födde i januari men speler inte fotboll (du ske visa att $A \cup B \setminus A \cap B \neq \emptyset$)

b) Mängden med grå färg = $\underbrace{A \cup B}_{\text{unionen av A och B}} \setminus \underbrace{A \cap B}_{\text{snittet av A och B}}$

minus

R /3

c) $P(A) = 0.3$, $P(A \cap B) = 0.1$

För att A och B ska vara oberoende

måste $P(A \cap B) = P(A) \cdot P(B)$

R

d.v.s $P(B) = \dfrac{P(A \cap B)}{P(A)} = \dfrac{0.1}{0.3} = \dfrac{1}{3} \approx 0.3334$

Svar: Sannolikheten för B måste vara ~33%

för att A och B ska vara oberoende.

/4

SU, STATISTIK

Skrivsal: ____________ Anonymkod: 0019-WTN Blad nr: 2

1)

d) R försöket är slumpmässigt och utfallsrummet är oändligt eftersom vi inte vet vad den "övre gränsen" är för studenternas vikt (i praktiken

R är det förstås ändligt eftersom ingen människa väger mer än 400 kg).

R Det är en kvotskala, eftersom det finns en absolut nollpunkt (0 kg) och en person som väger 90 kg kan sägas väga dubbelt så mycket som en som väger 45 kg.

/5

SU, STATISTIK

Skrivsal: ______ Anonymkod: 0019-WTN Blad nr: 3

2) A = "Blir antagen till Beckmans"

B = "Blir antagen till Konstfack"

$P(A) = 0.8$ $\quad P(A \cap B) = 0,3$

R $P(B) = 0,4$

a) vad är $P(\bar{A} \cap \bar{B})$?

R $P(\bar{A}) = 1 - P(A) = 1 - 0,8 = 0,2$

R $P(\bar{B}) = 1 - P(B) = 1 - 0,4 = 0,6$

$P(A) = P(A \cap B) + P(A \cap \bar{B})$ total sannolikhet

$\Rightarrow P(A \cap \bar{B}) = P(A) - P(A \cap B) = 0,8 - 0,3 = 0,5$

$P(\bar{B}) = P(A \cap \bar{B}) + P(\bar{A} \cap \bar{B})$ total sannolikhet

$\Rightarrow P(\bar{A} \cap \bar{B}) = P(\bar{B}) - P(A \cap \bar{B}) = 0,6 - 0.5 = \underline{\underline{0.1}}$

R

Svar: Det är 10% sannolikhet att Hanna inte erbjuds plats på varken Beckmans eller Konstfack.

/5

SU, STATISTIK

Skrivsal: ______ Anonymkod: 0019-WTN Blad nr: 4

2) b) Om X = "sannolikheten att bli inskriven på någon skola"
så är Y = "sannolikheten att <u>inte</u> bli inskriven på någon skola"

R $Y \sim Bin(6, 0.1)$ svaret från 2a = 10%

$P(X \geq 4) = P(Y \leq 2)$ dvs. sannolikheten

att minst 4 av 6 accepteras är samma

som att max 2 av 6 inte accepteras.

$$P(Y \leq 2) = P(Y=2) + P(Y=1) + P(Y=0)$$

$$\boxed{f(x) = \binom{n}{x} p^x (1-p)^{n-x}} \qquad \boxed{\binom{n}{x} = \frac{n!}{x!(n-x)!}}$$

$$\Rightarrow f(2) = \binom{6}{2} 0.1^2 \cdot 0.9^4 = 15 \cdot 0.006561 = 0.098415$$

R
$$f(1) = \binom{6}{1} 0.1^1 \cdot 0.9^5 = 6 \cdot 0.059049 = 0.354294$$

$$f(0) = \binom{6}{0} 0.1^0 \cdot 0.9^6 = 1 \cdot 0.531441 = 0.531441$$

$$f(2) + f(1) + f(0) = 0.098415 + 0.354294 + 0.531441 = \underline{\underline{0.98415}}$$

R

Svar: Det är alltså minst 98% sannolikhet att minst 4 av 6 kommer accepteras till minst en av skolorna.

/ III

SU, STATISTIK

Skrivsal: ____________ Anonymkod: 0019-WTN Blad nr: 5

2/ c) för första plats (5000 kr) kan du välja bland 90 personer

för andra plats (2000 kr) kan du nu bara välja bland 89 personer

för tredje plats (1000 kr) kan du bara välja bland 88 personer.

multiplikationsprincipen ger då att det

x finns totalt $90 \cdot 89 \cdot 88 = \underline{\underline{704880}}$

sätt att välja dessa tre personer.

/5

d/ för att anta att man kan approximera binomialfördelning till en normalfördelning

så måste både $n \cdot p > 5$ och $n(1-p) > 5$

$n \cdot p = 6 \cdot 0.9 = 5.4$ men $n(1-p) = 6 \cdot 0.1 = 0.6 < 5$

alltså kan vi <u>inte</u> anta normalfördelning för variabeln "antal antagna studenter".

/4

SU, STATISTIK

Skrivsal: ________ Anonymkod: 0019WTN Blad nr: 6

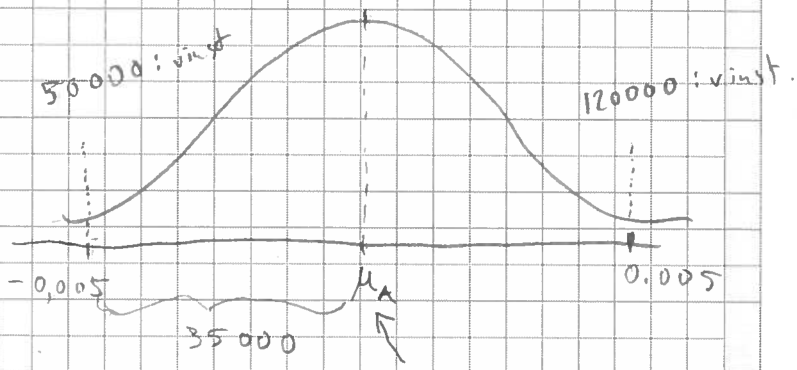
3) a) Vad är $E(x_A)$ eller μ_A?

$$P(X < 50000 \text{ kr vinst}) = 0.005$$

R $$P(X > 120000 \text{ kr vinst}) = 0.005$$

$$\Rightarrow P(X < 120000 \text{ kr vinst}) = 0.995$$

Eftersom resultatet är en normal fördelad stokastisk variabel så kommer μ_A ligga i mitten av kurvan



$$\frac{120000 - 50000}{2} = 35000$$

50000 + 35000 = 85000

$$\mu_A = 50000 + 35000 = \underline{\underline{85000}}$$

/4

Svar: företags A förväntade vinst är 85000 kr.

SU, STATISTIK

Skrivsal: ____________ Anonymkod: 0019-WTN Blad nr: 7

3/b)

Företag B: vänte värde 50000 kr och $\sigma = 100$

$$B \sim N(50000, 100^2)$$

Förväntat sammanlagt resultat = förväntat resultat för A + förväntat resultat för B = 85000 + 50000

$= \underline{135\,000 \text{ kr}}$

/3

c) Vad är σ_A? (standardavikelse för resultat för företag A)

$$P\left(\frac{X-\mu_A}{\sigma_A} \leq \frac{50000-\mu_A}{\sigma_A}\right) = \underbrace{0.005}_{\text{från tabell}}$$

när $\alpha = 0{,}005$

så är $Z_\alpha = 2{,}5758$

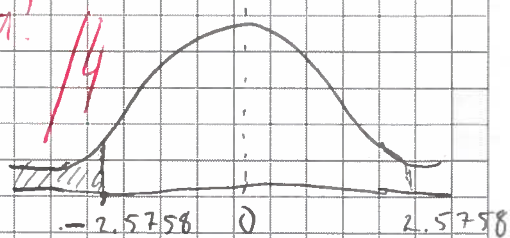
$$\Rightarrow Z = \frac{50000 - 85000}{\sigma_A}$$

p.g.a symmetri

$$-2.5758 = \frac{-35000}{\sigma_A} \qquad \Rightarrow Z_\alpha = -2{,}5758$$

$$\Rightarrow \sigma_A = \frac{-35000}{-2.5758} = 13588$$

bra! /4



Svar: Standardavikelsen för företag A är $\underline{13588}$ kr

SU, STATISTIK

Skrivsal: ______ Anonymkod: 0019-WTN Blad nr: 8

3/ c)

$$\text{Korrelation} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}}$$

Enligt uppgift far kovariansen mellan företagen 40000.

$$\text{Var}(A) = \sigma_A^2 = 13588^2$$

$$\text{Var}(B) = \sigma_B^2 = 100^2$$

Korrelationen: $\frac{40000}{\sqrt{(13588)^2 \cdot (100)^2}} = 0.029438$

Svar: Detta är en svagt positiv korrelation eftersom korrelation kan gå mellan −1 till +1, dvs. 0.029 är positiv men nära noll, alltså svag.

1/6

$\text{Var}(X+Y) = ?$

SU, STATISTIK

Skrivsal: ______ Anonymkod: 0019-WTN Blad nr: 9

4/ a)

X (vara A), Y (vara B)

Y \ X	0	1	2	3	$f_Y(y)$
1	0,25	0,25	0,15	0,1	0,75
2	0	0	0,25	0	0,25
$f_X(x)$	0,25	0,25	0,4	0,1	1

$f_{X,Y}(2,2) = 0,25$

R /5

b) $$P(X \mid Y) = \frac{P(X \cap Y)}{P(Y)}$$

fördelning av betingad efterfrågan

$$P(X=0 \mid Y=1) = \frac{0.25}{0,75} = \frac{1}{3} \approx 0.3333$$ R

$$P(X=1 \mid Y=1) = \frac{0.25}{0,75} = \frac{1}{3} \approx 0.3333$$ R

$$P(X=2 \mid Y=1) = \frac{0,15}{0,75} = 0,2$$ R

$$P(X=3 \mid Y=1) = \frac{0,1}{0,75} \approx 0.1333$$ R

/5

SU, STATISTIK

Skrivsal: ____________ Anonymkod: 0019-WTN Blad nr: 10

4/c)

Tabell för förväntad efterfrågan av vara A givet att 1 enhet av B efterfrågats

X	$f(x)$	$x \cdot f(x)$	$x^2 \cdot f(x)$
0	$\frac{1}{3}$	0	0
1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
2	0.2	0.4	0.8
3	0.1333	0.4	1.2
		$\underbrace{\sum 1.1333}_{E(x)}$ R	$\sum 2.3333$

$$Var(X) = EX^2 - (E(x))^2 = 2.3333 - (1.1333)^2$$

R $= \underline{1.04893}$

Svar: Givet att en enhet av vara B efterfrågats så är förväntad efterfrågan av vara A 1.1333 med en varians på 1.04893

/10

SU, STATISTIK

Skrivsal: ______ Anonymkod: 0019-WTN Blad nr: 11

5/ a/ Nej, funktionen i figur I kan inte vara en kumulativ fördelningsfunktion. Anledningen är att varje händelse X måste ha en sannolikhet att inträffa mellan $0 < p < 1$, d.v.s mellan noll och ett. Detta är ett av Kolmogorovs axiom.
Enligt figuren så sjunker sannolikheten vid $X = 1,5$, men det kan den bara göra om $p(X = 1,5) < 0$, vilket inte är tillåtet.

R bra /3

5 b) –

SU, STATISTIK

Skrivsal: ____________ Anonymkod: 0019-WTN Blad nr: 12

5/c)

$P(X \geq 7)$, $X \sim Bin(20, 0.2)$

test för att se om den kan approximeras till en normalfördelning:

$n \cdot p = 20 \times 0.2 = 4 < 5$, d.v.s går ej.

$P(X \geq 7) = 1 - P(X \leq 6)$

så vad är $1 - \underbrace{P(X \leq 6)}_{\text{från tabell}}$ då $X \in Bin(20, 0.2)$?

$\Rightarrow 1 - 0.9133\,1 = \underline{\underline{0,08669}}$ R

Svar: Sannolikheten att $X \geq 7$ är 8.7 %

1/4

SU, STATISTIK

Skrivsal: ______ Anonymkod: 0019-WTN Blad nr: 13

5/d)

$P(X \geq 36)$, $X \sim Bin(80, 0.4)$

test för att se om den kan approximeras till en normalfördelning:

R

$n \cdot p = 80 \cdot 0{,}4 = 32 > 5$

$n \cdot (1-p) = 80 \cdot 0{,}6 = 48 > 5$

ja, det går att approximera till en normalfördelning

R $P(X \geq 36) = 1 - P(X \leq 35)$, R

$\mu_x = n \cdot p = 80 \cdot 0.4 = 32$

$\sigma_x^2 = np(1-p) = 32(1-0.4) = 19.2$

$\Rightarrow 1 - P(X \leq 35) = 1 - P\left(\frac{X-\mu_x}{\sigma_x} < \frac{35{,}5 - 32}{\sqrt{19.2}}\right)$

kont. korrektion

$= 1 - P(\underbrace{Z < 0.684653}_{\text{från tabell}}) = 1 - 0.753 = \underline{\underline{0.247}}$

ett värde mellan

OK 0.75175 och $0.75490 \approx 0.753$

Svar: Sannolikheten för att $X \geq 36$ är 24.7%

/4.5

SU, STATISTIK

Skrivsal: ______ Anonymkod: 0019-WTN Blad nr: 14

5/e) $P(3 < X \le 6)$, $X \sim N(2, \sigma^2 = 9)$

dvs förväntat värde 2
och standardavvikelse 3

p.g.a. kontinuitets korrektion: N-fördelning är kont. från början!

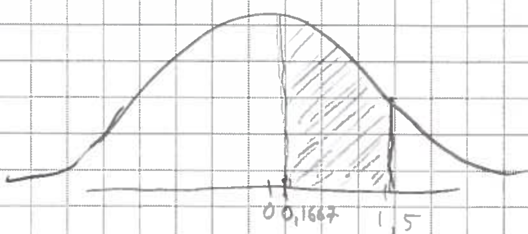
$$P(2,5 < X \le 6,5)$$

$$P\left(\frac{2,5-\mu}{\sigma} < \frac{X-\mu}{\sigma} \le \frac{6,5-\mu}{\sigma}\right)$$

$$P\left(\frac{2,5-2}{3} < Z \le \frac{6,5-2}{3}\right)$$

$$P\left(\frac{0,5}{3} < Z \le \frac{4,5}{3}\right)$$

$$P(0,16667 < Z \le 1,5)$$



$$P(Z > 0.16667) = 1 - P(Z < 0.16667)$$

$$P(Z \le 1,5) - (1 - P(Z < 0.16667))$$

$$\Phi(1,5) - 1 + \Phi(0,16667) = 0.93319 - 1 + 0,56749$$

$$= 0,50068$$

Svar: sannolikheten att $3 < X \le 6$ är 50 %

2-