**STOCKHOLMS UNIVERSITET**
Statistiska institutionen
Jolanta Pielaszkiewicz

# TENTAMEN I STATISTIKENS GRUNDER 1
## 2018-11-05

---

**Skrivtid:** 10.00-15.00
**Godkända hjälpmedel:** Miniräknare, språklexikon.
Tentamen består av fem uppgifter. För full poäng på en uppgift krävs tydliga, utförliga och väl motiverade lösningar.

---

**Uppgift 1.** (20 poäng)

a) Kan två händelser samtidigt vara beroende och disjunkta? Ge exempel och visa att det är möjligt eller motivera kort varför det är omöjligt.

b) I en fabrik tillverkas produkter i tre olika maskiner. Den första producerar 25 % av enheterna, den andra 30% av enheterna och den tredje 45% av enheterna. Maskinerna tillverkar en viss andel defekta enheter (5%, 3%, 1% är sannolikheten för defekt vid användning av de respektive maskinerna).

- Beräkna sannolikheten att en slumpmässigt vald enhet är defekt.
- En kund fick en felaktig enhet vid ett köp. Hur stor är sannolikheten att den har tillverkats i den andra maskinen?

**Uppgift 2.** (20 poäng)

a) På en tågstation finns tio olika spår som är tillgängliga för kommande tåg med sannolikhet 80%. Spårens tillgänglighet är oberoende av varandra. Vad är sannolikhet att minst 9 av 10 spår är tillgängliga?

b) Järnvägsnätet i Sveige är indelat i 96 numrerade stråk. Sannolikheten att ett elfel inträffar på något stråk är 7%.
Beräkna sannolikhen att elfelet inträffar på minst 10 men inte mer än 40 stråk i järnvägsnätet.

c) Låt $X$ =*antal tillgängliga spår* (se deluppgift 2a)) och $Y$ =*antal stråk med elfel* (se deluppgift 2b)). Anta att $X$ och $Y$ är oberoende.
Beräkna sannolikheterna och tolka resultatet.

c1) $P(X = 5, Y = 0) = f_{X,Y}(5, 0)$
c2) $P(X = 5|Y = 0) = f_{X|Y}(5|0)$

**Uppgift 3.** (20 poäng) Anna har bjudit in 10 kollegor på middag kl. 19 på fredag. Hon får sms från okänd telefonnummer att 2 inbjudna personer blir sjuka och kan tyvärr inte vara med.

a) Hur många olika sällskap är möjliga på middag (vi vet att bara 8 av 10 gästerna kommer)?

b) Vi vet att gästens ankomsttid är normalfördelad med väntevärdet 19 och standardavvikelse 10 minuter. Vad är sannolikheten att alla 8 personer kommer innan kl. 19:10?
Anta att gästens ankomsttiderna är oberoende av varandra.

**Uppgift 4.** (20 poäng) Låt oss ha en undersökning med två oberoende stokastiska variabeler:

- $X$ som är normalfördelad med väntevärde 6 och standardavvikelse lika med 2, dvs. $X \sim N(6, \sigma^2 = 4)$
- $W$ som är Bin(4.0.5)-fördelad.

Dessutom vet vi att $X$ och $W$ är oberoende.

a) Beräkna variansen av $\frac{1}{3}X + W - 1$.

b) Låt $Y = 2X - W$. Beräkna kovariansen $Cov(X, Y)$.

c) Rita frekvensfunktionen $f_W(w)$ och kumulativa fördelningsfunktionen $F_W(x)$ för $W$.

d) Låt oss ha stokastisk variabln $Z$ som har Bernouli fördelning med p=0.7. Fyll i tabellen nedan med den simultana sannolikhetsfördelningen för variablerna $W$ och $Z$ och beräkna $Cor(W, Z)$. Är $W$ och $Z$ oberoende?

| | | W: 0 | 1 | 2 | 3 | 4 | $f_Z(z)$ |
|---|---|---|---|---|---|---|---|
| | 0 | | 0.1 | | 0.1 | | 0.3 |
| Z | 1 | 0.05 | | 0.3 | | 0.05 | 0.7 |
| $f_W(w)$ | | | | | | | |

**Uppgift 5.** (20 poäng) $X$ är normalfördelad med $\mu = 100$ och $\sigma = 5$.

a) Beräkna $P(X \leq 110)$.

b) Beräkna $P(30 \leq X \leq 110)$.

c) Beräkna $P(X \geq \frac{1}{2}X)$.

d) Beräkna $P(-2 \leq Y < 0)$, där $Y \sim \chi^2(10)$-fördelad.

e) Stokastisk variabeln $T$ är t-fördelad. Hitta antalet frihetsgrader så att $P(T > 2) \approx 0.05$.

f) Kan man beräkna $P(X \geq \mu)$ om väntevärde $\mu$ är okänd (ej 100 som tidigare)? Beräkna denna sannolikhet om det är möjligt eller motivera kort varför det inte går att göra.

Statistiska institutionen

Stockholms universitet

# Rättningsblad

**Datum:** 5/11/18

**Sal:** Ugglevikssalen

**Tenta:** Statistikens grunder

**Kurs:** Statistikens grunder 1

**ANONYMKOD:** 0029-HGL

☑ 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 | | | | | 9st |
| Lär.ant. 10 | 15 | 18 | 16 | 12 | | | | | |

GC

| POÄNG | BETYG | Lärarens sign. |
|---|---|---|
| 71 | C | [illegible] |

**SU, STATISTIK**

Skrivsal: Ugglevikssalen Anonymkod: 0029-HGL Blad nr: 1

Uppgift 1

Disjunkta kan skrivas som $P(A) = P(B)$ alltså att händelse A ej är samma som händelse B. A - händelse att man är äldre än 18 B - händelse att man är äldre än 20 ej samma ej disjunkta

R Om beroende gäller, gäller formeln $P(A) \cdot P(B) \neq P(A \cap B)$

Vilket ger $P(A) \cdot P(B) = P(A \cap B)$ om de är disjunkta. Alltså kan de R vara disjunkta men även beroende. ?

Det betyder att A är olik B vilket ger A eller B. /2

b) Maskin A producerar 0,25 och ger defekta enheter med 0,05

Maskin B producerar 0,3 och ger defekta enheter med 0,03

Maskin C producerar 0,45 och ger defekta enheter med 0,01

Om vi benämner D = defekt söker vi $P(A|D) + P(B|D) + P(C|D)$

Vilket ger $(0{,}25 \cdot 0{,}05) + (0{,}3 \cdot 0{,}03) + (0{,}45 \cdot 0{,}01) = 0{,}026$ /8 −

Som är avrundat uppåt ≈ 3%

Nu söker vi sannolikheten att en enhet har tillverkats i den andra maskinen "B", givet att vi fått en felprodukt

$$P(D|B) = \frac{P(B|D) \cdot P(D)}{\sum_{i=1} P(A|D) \cdot P(D) + P(B|D) \cdot P(D) + P(C|D) \cdot P(C)} = \frac{(0{,}3 \cdot 0{,}03)(0{,}03)}{0{,}00094} = \frac{0{,}00027}{0{,}00094} = 0{,}287$$

≈ 29%

Vi använde oss av Bayes sats för att räkna ut sannolikheten

**SU, STATISTIK**

Skrivsal: Ugglevikssalen Anonymkod: 0029-HGL Blad nr: 2

uppgift 2

a Vi vet att sannolikheten att ett tåg kommer in på ett spår är 0,8.

Vi vet att det finns 10 spår som är oberoende och vi söker

$P(X \geq 9)$ som kan skrivas $1 - P(X \leq 8)$

$f(9) + f(10)$ R

$f(0) + f(1) + f(2) + f(3) + f(4) + f(5) + f(6) + f(7) + f(8)$ R

Vi kan nu se att X är binomialfördelad så att $X \sim Bin(n=10, p=0,8)$

$f(9) = \binom{10}{9} \cdot 0,8^9 \, 0,2^1 = \frac{10!}{9!(1)!} \cdot 0,8^9 \, 0,2^1 = 10 \cdot 0,8^9 \cdot 0,2^1 = 0,2684$ R

$+ f(10) = \frac{10!}{10!(0)!} \cdot 0,8^{10} \cdot 0,2^0 = 1 \cdot 0,8^{10} \cdot 1 = 0,10737$ R som sammanlagt ger

$\approx \underline{0,3758}$ som är cirka 38% R

/5

b Vi vet att vi har 96 stråk och "p" = 0,07 att ett elfel inträffar.

Vi vet att X är binomialfördelad $X \sim Bin(n=96, p=0,07)$

Dock ser vi att $n \cdot p = 96 \cdot 0,07 = 6,72$ och $n \cdot (1-p) = 96 \cdot 0,93 = 89,28$

kan vi normalapproximera fördelningen. Vi behöver först använda oss av kontinuitetskorrektion R

$\Rightarrow P\left(\frac{x - \frac{1}{2} - np}{\sqrt{n \cdot p \cdot (1-p)}} < \mu < \frac{x + \frac{1}{2} - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}}\right)$ R

eftersom binomialfördelningen endast hanterar diskreta tal för medan normalfördelningen räknar med kontinuerliga.

vänd →

SU, STATISTIK

Skrivsal: Ugglevikssalen Anonymkod: 0029 - HGL Blad nr: 3

Uppgift 2 fortsättn.

$P(10{,}5 \leq X \leq 40)$

Vi har nu förklarat lite regler. Vi söker $P(X \geq 10 / Y \leq 40)$

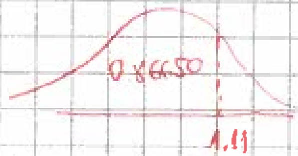
Vi benämner den andra sannolikheten med "Y."

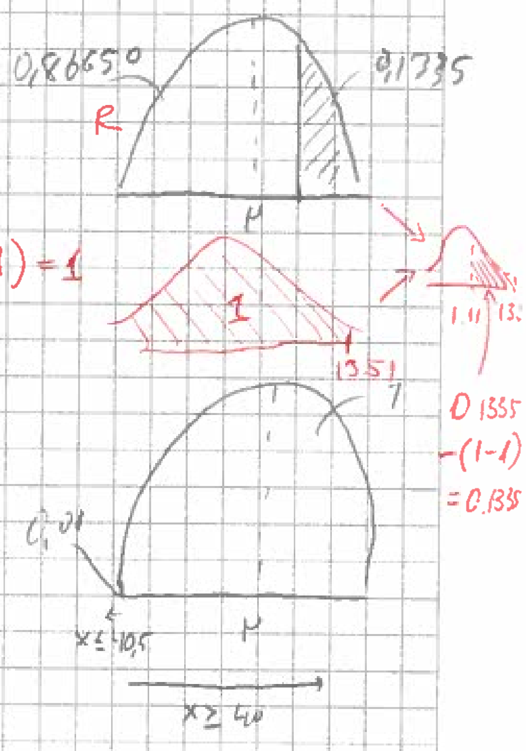
Vi vet att $E(X) = n \cdot p = 6{,}72$ och $V(X) = n \cdot p \cdot (1-p) = 6{,}2496$

För X räknar vi $P(X \geq 10)$

$$\Rightarrow P\left(X \geq \frac{10 - 0{,}5 - 6{,}72}{\sqrt{6{,}2496}}\right) \underset{\text{approx } N(0,1)}{\approx} \Rightarrow P(Z \geq \phi\, 1{,}11) \text{ vilket ger } 1 - 0{,}86650 = 0{,}1335$$

$\phi(1{,}11) = 0{,}86650$




Nu räknar vi för "Y" $P(Y \leq 40)$

$P(Z \leq 13{,}51) = \phi(13{,}51) = 1$

$$P\left(X \leq \frac{40 + 0{,}5 - 6{,}72}{\sqrt{6{,}2496}}\right) \underset{\text{approx } N(0,1)}{\approx} \Rightarrow P(Z \leq \phi\, 13{,}51)$$

$\approx 0$

När vi räknat $P(X \geq 10 / Y \leq 40)$ måste vi multiplicera sannolikheterna vilket ger

$0{,}1335 \cdot 0 = 0$ Det är alltså cirka 0% sannolikhet att detta inträffar.

/7

vänd →

**SU, STATISTIK**

**Skrivsal:** Uggleviksssalen **Anonymkod:** 0029-HGL **Blad nr:** 4

Uppgift 2 fortsättar

c1 Vi vet att $X$ = antal spår = 10 och $Y$ = Stråk med elfel = 7% av 96.

$P(X=5, Y=0)$ ger oss $X \sim Bin(n=10, p=0{,}8)$ och vi söker $P(X=5)$

Som ger $f(5) \Rightarrow \binom{10}{5} 0{,}8^5 \cdot 0{,}2^5 = \frac{10!}{5!(5)!} \cdot 0{,}8^5 \cdot 0{,}2^5 = 252 \cdot 0{,}8^5 \cdot 0{,}2^5$
$= 0{,}02642$ R /1,5

Och $P(Y=0)$ ger $X \sim Bin(n=96, p=0{,}07)$ men där $n \cdot p > 5$ och $n \cdot (1-p) > 5$

$E(X) = n \cdot p = 6{,}72$ $V(X) = n \cdot p(1-p) = 6{,}2496$ Så vi normalapproximerar

$P(Y=0) \Rightarrow P\left(\frac{0-0{,}5-6{,}72}{\sqrt{6{,}2496}} \overset{\text{approx}}{\sim} N(0,1) < Y < \frac{0+0{,}5-6{,}72}{\sqrt{6{,}2490}} \overset{\text{approx}}{\sim} N(0,1)\right)$

$P(-2{,}888 < Z < -2{,}488)$

$\Rightarrow P(\Phi -2{,}9 < Z < \Phi -2{,}48)$ R /1,5

$Z = 0{,}00657 - 0{,}00187 = 0{,}0047$ Vi adderar nu $0{,}02642 + 0{,}0047$

$\Phi(-2{,}488) - \Phi(-2{,}888) = 1 - \Phi(2{,}488)$ $= 0{,}03112 \approx 3\%$ "$P(X=5, Y=0)$"

$- 1 + \Phi(2{,}888) = \Phi(2{,}888) - \Phi(2{,}488) = \cancel{0{,}00199}$
$0{,}99807 - 0{,}99361 = 0{,}00446$

c2 Vi söker $P(X=5 / Y=0)$ Vi använder samma siffror från föregående uppgift.

Dock säger "$f_{X/Y}$" $\neq f(x) \cdot f(y)$ som ger $0{,}02642 \cdot 0{,}0047 = 0{,}000124$

som blir cirka 0.

–

/3p

**SU, STATISTIK**

Skrivsal: Ugglevikssalen Anonymkod: 0029-HGL Blad nr: 5

Uppgift 3 Vi vet att 2 av 10 har uteslutits och vi har därför endast 8 personer kvar

Hur många möjliga permutationer kan vi få av 8 personer?
Vi antar att plats spelar roll. ← Varför? Det är viktig vem kommer inte var någonstans sitter han eller?

Vi får då $8! = 1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7\cdot8 = 40320$ olika permutationer av 8 personer

↑ antal av sittningar av 8 gäster

/5

b) Vi vet att ankomsttid är normalfördelad, vi betecknar ankomsttid som $X \sim N(0,1)$

Där $\mu = 19$ och $\sigma_x = 10$ minuter. Vi söker $P(X < 19{,}10)$

$P\left(X < \frac{19.10 - 19}{0.10}\right) = P(Z < 0.01)$ 0,50399

/8

Vi vill nu veta sannolikheten för att detta händer alla 8 personer.
Vi får då $0{,}50399^8 = 0{,}00416 \approx 0{,}42\,\%$

/5

Det är alltså inte så stor sannolikhet.

**SU, STATISTIK**

Skrivsal: Ugglevikssalen  Anonymkod: 0029-HGL  Blad nr: 6

Uppgift 4  Vi har 2 oberoende variabler, X och W

$X \sim N(6, 2^2)$  $W \sim Bin(n=4, p=0,5)$

De antas oberoende

a  Vi vet att $\mu_X = 6$ och $V(X) = 4$  Variansen av X om $\frac{1}{3}X$

så är $V(X) = (\frac{1}{3})^2 \cdot X \Rightarrow \frac{1}{9} \cdot X = 0,111 \cdot 4 = 0,444$

W, vi vet att E(W) för $W \sim Bin(n,p)$ är $n \cdot p$ som blir $4 \cdot 0,5 = 2$

och V(W) blir $n \cdot p \cdot (1-p) = 4 \cdot 0,5 \cdot 0,5 = 1$

Då blir $V(\frac{1}{3}X + W - 7) = 0,444 + 1 - 7 = 0,444$ eller $4(\frac{1}{3})^2$  /3

b  Om $Y = 2X - W$ blir $Cov_{X,Y}$ Var (Y) $= 2 \cdot a^2 \cdot V(X) + b^2 \cdot V(W) - 0$ ← $2ab\,Cov_{XY}$ faller, med Eftersom de är oberoende

Som ger $(2 \cdot (\frac{1}{3})^2 \cdot 4) - (1^2 \cdot n \cdot p \cdot (1-p)) = 0,888 - 1 = -0,111$

Men eftersom de är oberoende har de Covariansens värde = 0

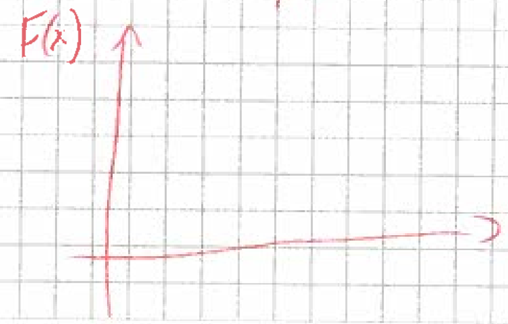
Alltså har de heller ingen korrelation  /1

c  för $W \sim Bin(n=4, p=0,5)$

Kumulativ fördelning blir

| f(x) X | P | F(x) | $X^2$ | F(x) |
|---|---|---|---|---|
| 0 | 0,5 | 0,0625 | 0,0039 | 0,0625 |
| 1 | 0,5 | 0,25 | 0,0625 | 0,3125 |
| 2 | 0,5 | 0,375 | 0,140625 | 0,6875 |
| 3 | 0,5 | 0,25 | 0,0625 | 0,9375 |
| 4 | 0,5 | 0,0625 | 0,0039 | 1 |
| 5 | | R | | R |

/4


värd →

**SU, STATISTIK**

Skrivsal: Ugglevikssalen Anonymkod: 0029-HGL Blad nr: 7

Uppgift 4 fortsätter

OBS! Siffror inom parentes är redan uträknade

Vi vet att $Z \sim$ Bernoulli med $p = 0,7$ $\begin{cases} 0 = 0,3 \\ 1 = 0,7 \end{cases}$

| Z \ W | 0 | 1 | 2 | 3 | 4 | |
|---|---|---|---|---|---|---|
| 0 | 0,0125 | (0,1) | 0,075 | (0,1) | 0,0125 | ⇒ 0,3 |
| 1 | (0,05) | 0,15 | (0,3) | 0,15 | (0,05) | ⇒ 0,7 |
| R | 0,0625 | 0,25 | 0,375 | 0,25 | 0,0625 | 1 |

$W(0) = \frac{4!}{0!(4)!} \cdot 0,5^0 \cdot 0,5^4 = 0,0625$

$W(1) = \frac{4!}{1!(3)!} \cdot 0,5^1 \cdot 0,5^3 = 0,25$

$W(2) = \frac{4!}{2!(2)!} \cdot 0,5^2 \cdot 0,5^2 = 0,375$

$W(3) = \frac{4!}{3!(1)!} \cdot 0,5^3 \cdot 0,5^1 = 0,25$

$W(4) = \frac{4!}{4!(0)!} \cdot 0,5^4 \cdot 0,5^0 = 0,0625$

$\Sigma = 1$ R

$E(X_W) = (0 \cdot 0,0625) + (1 \cdot 0,25) + (2 \cdot 0,375) + (3 \cdot 0,25) + (4 \cdot 0,0625) = \underline{2}$

$V(X_W) = (0-2)^2 \cdot 0,0625 + (1-2)^2 \cdot 0,25 + (2-2)^2 \cdot 0,375 + (3-2)^2 \cdot 0,25 + (4-2)^2 \cdot 0,0625 = \underline{1}$ R

$E(X_Z) = (0 \cdot 0,3) + (1 \cdot 0,7) = \underline{0,7}$ R

$V(X_Z) = (0-0,7)^2 \cdot 0,3 + (1-0,7)^2 \cdot 0,7 = \underline{0,21}$

$Cov(W,Z) = (0-2) \cdot (0-0,7) \cdot 0,0125 + (1-2) \cdot (0-0,7) \cdot 0,1 + (2-2) \cdot (0-0,7) \cdot 0,075 + (3-2) \cdot (0-0,7) \cdot 0,1 + (4-2) \cdot (0-0,7) \cdot 0,0125 + (0-2) \cdot (1-0,7) \cdot 0,05 + (1-2) \cdot (1-0,7) \cdot 0,15 + (2-2) \cdot (1-0,7) \cdot 0,3 + (3-2) \cdot (1-0,7) \cdot 0,15 + (4-2)(1-0,7) \cdot 0,05$

$= 0,0175 + 0,07 + 0 - 0,07 - 0,0175 - 0,03 - 0,045 + 0 + 0,045 + 0,03 = \underline{0}$

ber.

ofm Vi kan även se att om beroende sker så är $P(A) \cdot P(B) = P(A \cap B)$

Test: för $W(0)$ och $Z(0)$ ⇒ $0,0625 \cdot 0,3 = 0,01875 \neq 0,0125$ !

/8

**SU, STATISTIK**

Skrivsal: Uggleviksssalen Anonymkod: 0029-HGL Blad nr: 8

Uppgift 5 vi vet att $X \sim N(100, 5^2)$

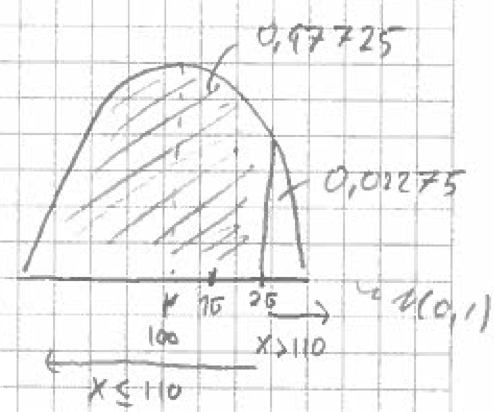
a vi söker nu $P(X \leq 110)$ vilket ger

$\Rightarrow P\left(\frac{X-\mu}{\sigma} \leq \frac{110-100}{5}\right) \Rightarrow P(Z \leq \phi 2) = \phi(2) =$

$\underbrace{\phantom{Z}}_{N(0,1)}$ 0,97725

R

1/1




b $P(30 \leq X \leq 110) \Rightarrow P\left(\frac{30-100}{5} \leq \frac{X-\mu}{5} \leq \frac{110-100}{5}\right)$ ← $N(0,1)$

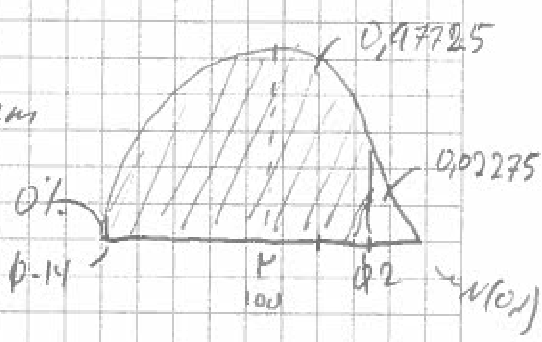
Notation

$\Rightarrow P(\phi\text{-}14 \leq Z \leq \phi 2)$ $\approx$ blir $\phi(2) - \phi(-14) = 0{,}97725 - 0$ R

0

Eftersom Sannolikheten att variabeln är mer extrem än −14 standardavvikelser är lika med 0

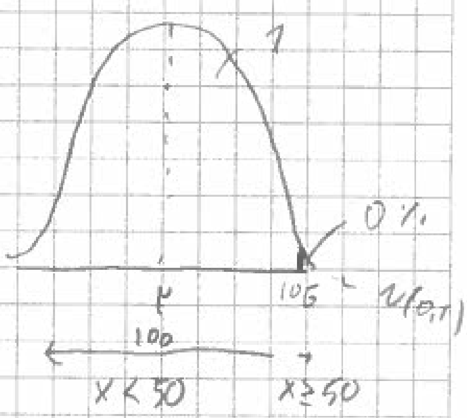
1/1




c $P(X \geq \frac{1}{2}X) \Rightarrow P\left(\frac{X-\mu}{\sigma} \geq \frac{100 \cdot 0{,}5}{5}\right)$ ← $N(0,1)$

$\Rightarrow P(Z \geq \phi 10)$ vilket ger $1 - \phi 10 = 0$

1/0




vänd →

**SU, STATISTIK**

Skrivsal: Ugglevikssalen  Anonymkod: 0029-HGL  Blad nr: 9

Uppgift 5 fortsätter

d vi söker $P(-21 \le Y < 0)$ där $Y = X^2(10)$ dgf

Vi kan normalapproximera då "μ" och "σ" kommer från en normalfördelning

$\Rightarrow P(-2 \le Y < 0) \Rightarrow$ vi vet att med $X^2$ så kvadrerar vi uttrycket

till $P(\frac{-2}{5} \le Y < \frac{0}{5})^2 \Rightarrow P(\frac{4-100}{25} \le Z < \frac{0-100}{25}) \Rightarrow P(-3{,}84 \le Z < -4)$ $\sim N(0,1)$

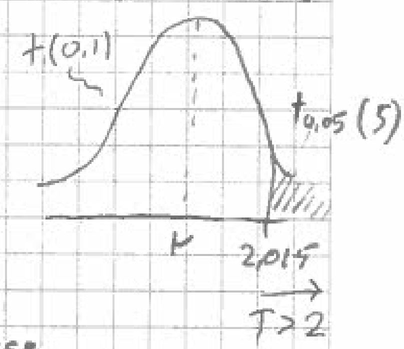
$\Rightarrow 0 - 0 \approx 0$ Sannolikheten är alltså cirka 0.

Y är inte negativt om $\chi^2$-fördelning  /0

e vi söker $P(T > 2) = 0{,}05$ alltså har vi signifikansnivå på $\alpha = 0{,}05$

vi vet att $t_{0,05}(x) > 2$ där $t_{0,05}(5) = 2{,}015$ dgf

vilket är större än 2.  /3




f Det går att räkna ut $P(X \ge \mu)$ så länge vi vet sannolikheten, har vi sannolikheten kan vi se hur många standardavvikelser som avviker från vår hypotetiska urvalsmedelvärde $\bar{x}$.  /1

Ex
vi vet vår $\bar{x}$ som är 10 och $\sigma = 2$ och $p = 0{,}01$

$\Rightarrow P(\frac{\mu - 10}{2}) = 0{,}01$ vi vet att 0,01 är -2,3263 standardavvikelser från väntevärdet.

$\Rightarrow P(\frac{\mu - 10}{2}) = -2{,}3263 \Rightarrow P(\mu - 10) = -4{,}6526$ $+10$  $\mu \approx 5{,}3474$

Men har vi inga parametrar blir det svårt!