STOCKHOLMS UNIVERSITET
Statistiska institutionen
Ellinor Fackle-Fornius

TENTAMEN I STATISTISK TEORI MED TILLÄMPNINGAR I

2017-09-26

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.
Resultatet meddelas senast den 12 oktober.

Uppgift 1. (20 poäng)

Basketproffset Bella Bollé avslutar alltid sina träningar med att öva på straffkast. Antag att hennes kast på träning är oberoende och att hon sätter 78 % av sina försök.

a) Antag att hon bestämmer sig för att göra 10 kast. Vad är sannolikheten att hon sätter minst 8 av dessa?

b) Antag istället att hon bestämmer sig för att gå hem när hon har satt 8 kast. Vad är sannolikheten att det krävs högst 10 kast innan hon får gå hem?

c) Bella skulle vilja öva upp sin träffprocent så att sannolikheten att hon sätter 10 av 10 kast är minst 0.9. Vilken träffprocent motsvarar det?

Uppgift 2. (20 poäng)

Antag att X och Y är två stokastiska variabler med täthetsfunktionerna

$$f(x) = \begin{cases} \frac{x}{2}, & 0 \leq x \leq 2 \\ 0, & \text{annars.} \end{cases}$$

och

$$f(y) = \begin{cases} y, & 0 < y < 1 \\ 2 - y, & 1 \leq y < 2 \\ 0, & \text{annars.} \end{cases}$$

a) Bestäm fördelningsfunktionen för X.

b) Bestäm fördelningsfunktionen för Y.

c) Beräkna väntevärdet för X.

d) Beräkna väntevärdet för Y.

Uppgift 3. (20 poäng)

Vid tillverkningen av flingpaket är påfyllningen av paketen ett viktigt moment. Helst ska endast en liten andel av paketen ha en nettovikt som understiger den påstådda nettovikten på 500 gram.

a) Antag att nettovikten av innehållet i flingpaketen är normalfördelad med väntevärde $\mu =$ 503 gram. Bestäm standardavvikelsen (σ) så att sannolikheten att nettovikten understiger 500 gram är 0.005.

b) Beräkna sannolikheten att nettovikten för ett flingpaket är mellan 502 och 504 gram. Använd μ och σ från a)-uppgiften.

c) Flingpaketen packas i lådor med 10 paket i varje låda. Vad är väntevärdet och standardavvikelsen för medelnettovikten för paketen i lådan? Använd μ och σ från a)-uppgiften samt antag oberoende mellan innehållet i paketen i lådan.

d) Beräkna sannolikheten att medelnettovikten för paketen i en låda (enligt c)-uppgiften) är mellan 502 och 504 gram.

Uppgift 4. (20 poäng)

Antag att hastigheten (Y) hos en gasmolekyl är en stokastisk variabel som följer Rayleigh-fördelningen. Denna fördelning har täthetsfunktionen

$$f(y) = \begin{cases} 2ye^{-y^2}, & y \geq 0 \\ 0, & \text{annars.} \end{cases}$$

Den kinetiska energin, U, hos en sådan gasmolekyl är en funktion av hastigheten enligt $U = 6Y^2$.

a) Bestäm täthetsfunktionen och fördelningsfunktionen för den kinetiska energin.

b) Vad kallas fördelningen för den kinetiska energin?

Uppgift 5. (20 poäng)

Företaget F^3 är en kedja av servicebutiker som bland annat erbjuder paketutlämning. Man är intresserad av att analysera de stokastiska variablerna Y_1, den totala tiden från att kunden ställer sig i kö för att göra ett paketärende till dess att ärendet är avslutat och Y_2, som är kötiden. Observera att $Y_2 \leq Y_1$. Antag att den simultana täthetsfunktionen för Y_1 och Y_2 i minuter ges av

$$f(y_1, y_2) = \begin{cases} e^{-y_1}, & 0 < y_2 \leq y_1 < \infty \\ 0, & \text{annars.} \end{cases}$$

a) Bestäm marginalfördelningarna för Y_1 och Y_2.

b) Är Y_1 och Y_2 stokastiskt oberoende?

c) Beräkna sannolikheten att den totala tiden (från att kunden ställer sig i kö för att göra ett paketärende till dess att ärendet är avslutat) understiger 2 min samtidigt som kötiden överstiger 1 minut.

Statistiska institutionen

Stockholms universitet

Rättningsblad

Datum: 26/09/17

Sal: Ugglevikssalen

Tenta: Statistikens teori med tillämpningar I

Kurs: Statistikens teori med tillämpningar I

ANONYMKOD: ST-0001

☒ 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					6
Lär.ant. 20	20	17	20	20					

97 + 5 Bonus

POÄNG	BETYG	Lärarens sign.
102	A	[signature]

SU, STATISTIK

Skrivsal: Vägleviк Anonymkod: ST-0001 Blad nr: 1

1 a) Y = antal kast hon sätter

$Y \sim bin(n=10,\ p=0{,}78)$ $y = 0, 1, \ldots, 10$ R

$P(Y \geq 8) = P(Y=8) + P(Y=9) + P(Y=10)$

$P(y) = \binom{n}{y} p^y (1-p)^{n-y}$ R

$P(Y \geq 8) = \binom{10}{8} 0{,}78^8 \cdot 0{,}22^2 + \binom{10}{9} 0{,}78^9 \cdot 0{,}22^1$

$+ \binom{10}{10} 0{,}78^{10} = 45 \cdot 0{,}78^8 \cdot 0{,}22^2 +$

$+ 10 \cdot 0{,}78^9 \cdot 0{,}22 + 1 \cdot 0{,}78^{10} \approx 0{,}6169$ R

Svar: Av 10 kast är sln ca 0,6169 att hon sätter minst 8 (dvs 8, 9 eller 10) av dessa kast 8

b) Y = antal kast tills hon har satt 8 kast $y = 8, 9, 10 \ldots$ R

$Y \sim Negbin(r=8,\ p=0{,}78)$ R

$P(Y \leq 10) = P(Y=8) + P(Y=9) + P(Y=10)$

$P(y) = \binom{y-1}{r-1} p^r (1-p)^{y-r}$ R

$P(Y \leq 10) = \binom{8-1}{8-1} 0{,}78^8 + \binom{9-1}{8-1} 0{,}78^8 \cdot 0{,}22^1 +$

$+ \binom{10-1}{8-1} 0{,}78^8 \cdot 0{,}22^2 = 1 \cdot 0{,}78^8 + 8 \cdot 0{,}78^8 \cdot 0{,}22 +$

$+ 36 \cdot 0{,}78^8 \cdot 0{,}22^2 \approx 0{,}6169$ R

Svar: Slh att det krävs högst 10 (dvs 8, 9 eller 10) kast innan hon har satt 8 kast och fä gå hem är ca 0,6169, dvs samma slh som att sätta minst 8 kast av 10 kast (som efterfrågades i a,). R 8

c, Y = antal kast hon sätter
$Y \sim bin(n=10, p=?)$ R
Vad ska p, dvs träffprocenten, vara för att $P(Y=10) \geq 0,9$?

$$P(y) = \binom{n}{y} p^y (1-p)^{n-y}$$

$$P(Y=10) = \binom{10}{10} p^{10} (1-p)^0 \geq 0,9$$ R

$$P(Y=10) = 1 \cdot p^{10} \cdot 1 = p^{10} \geq 0,9$$

$$p^{10} \geq 0,9 \Rightarrow p \geq \sqrt[10]{0,9}$$

$$p \geq 0,9896$$ R

Svar: För att slh att hon sätter 10 av 10 kast ska vara minst 0,9 måste p, dvs träffprocenten, vara minst 0,9896

4

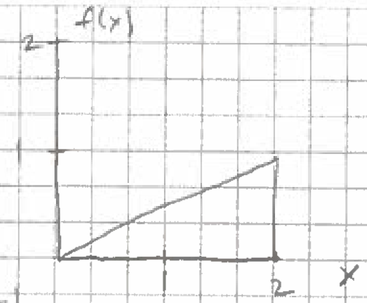
20

SU, STATISTIK

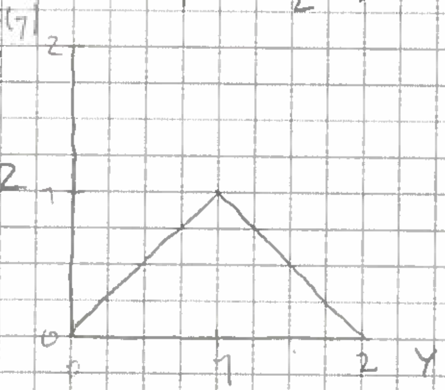
Skrivsal: Ugglevik **Anonymkod:** ST-0001 **Blad nr:** 2

(2)

$$f(x) = \begin{cases} \frac{x}{2}, & 0 < x < 2 \\ 0, & \text{annars} \end{cases}$$



$$f(y) = \begin{cases} y, & 0 < y < 1 \\ 2-y, & 1 < y < 2 \\ 0, & \text{annars} \end{cases}$$



a) $F(x) = \int_0^x \frac{x}{2}\,dx = \left[\frac{x^2}{4}\right]_0^x = \frac{x^2}{4}$ R

Svar:

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{4}, & 0 \leq x \leq 2 \\ 1, & x > 2 \end{cases}$$

R

4

b) Då täthetsfunktionen ser olika ut i olika intervall måste vi integrera de två intervallen separat och för det andra intervallet lägga på arean för det första intervallet

Bra!

$F(0<y<1) = \int_0^y y\,dy = \left[\frac{y^2}{2}\right]_0^y = \frac{y^2}{2}$ R

$F(1<y<2) = \int_1^y 2-y\,dy + \int_0^1 y\,dy =$

$= \left[2y-\frac{y^2}{2}\right]_1^y + \left[\frac{y^2}{2}\right]_0^1 = \left(2y-\frac{y^2}{2}\right)-\left(2-\frac{1}{2}\right)+\frac{1}{2} =$

$= 2y-\frac{y^2}{2}-2+\frac{1}{2}+\frac{1}{2} = 2y-\frac{y^2}{2}-1 = \frac{4y-y^2-2}{2}$ R

Svar:

$$F(y) = \begin{cases} 0, & y<0 \\ \frac{y^2}{2}, & 0<y<1 \\ \frac{4y-y^2-2}{2}, & 1\le y<2 \\ 1, & y>2 \end{cases}$$

R

6

c) $E(X) = \int_0^2 x\frac{x}{2}\,dx = \int_0^2 \frac{x^2}{2}\,dx = \left[\frac{x^3}{6}\right]_0^2 = \frac{2^3}{6} = \frac{4}{3}$

R 4

Svar Väntevärdet för X är $\frac{4}{3}$

d) Samma tänk här som uppgift 6, dvs då vi har olika täthetsfunktioner i olika intervall kommer intervallen behöva integreras separat och adderas. $\left(E(0<Y<2) = E(0<Y<1) + E(1<Y<2)\right)$ R

SU, STATISTIK

Skrivsal: Ugglevik **Anonymkod:** ST-0001 **Blad nr:** 3

(2) d, Forts.

$$E(Y) = \int_0^1 y \cdot y \, dy + \int_1^2 y \cdot (2-y) \, dy =$$ R

$$= \int_0^1 y^2 dy + \int_1^2 2y - y^2 \, dy = \left[\frac{y^3}{3}\right]_0^1 + \left[\frac{2y^2}{2} - \frac{y^3}{3}\right]_1^2 =$$

$$= \frac{1}{3} + \left(\frac{2 \cdot 2^2}{2} - \frac{2^3}{3}\right) - \left(\frac{2 \cdot 1^2}{2} - \frac{1^3}{3}\right) = \frac{1}{3} + 4 - \frac{8}{3} - 1 + \frac{1}{3} = 1$$ R

Svar: Väntevärdet för Y är 1 6

(20)

(3) a, Y = nettovikt av innehåll i flingpaket (g)

$Y \sim N(\mu = 503, \sigma)$

$P(Y < 500) = 0{,}005$ $P\left(Z < \frac{500-503}{\sigma}\right) = 0{,}005$

Enligt tabell: $P(Z < 2{,}58) \approx 0{,}995$

$\Rightarrow P(Z < -2{,}58) \approx 0{,}005 \Rightarrow$

$\frac{500-503}{\sigma} = -2{,}58$ $500 - 503 = -2{,}58\sigma$

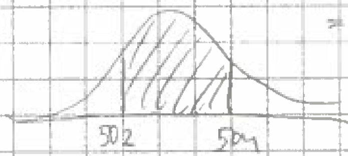
$\sigma = \frac{500-503}{-2{,}58}$ $\sigma = \frac{50}{43} \approx 1{,}1628$ R

Svar: Då sannolikheten att vikten är mindre än 500g är 0,005 är $\sigma = \frac{50}{43} \approx 1{,}1628$

5

$Y \sim N(\mu = 503, \sigma = 50/43)$

b) $P(502 < Y < 504) =$



$= P(Y < 504) -$

$- P(Y < 502) = P\left(Z < \frac{504-503}{50/43}\right) - P\left(Z < \frac{502-503}{50/43}\right) =$

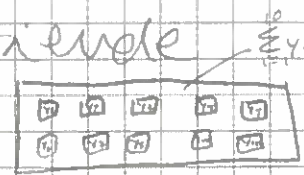
$= P(Z < 0{,}86) - P(Z < -0{,}86) = \Phi(0{,}86) - (1 - \Phi(0{,}86))$

$= 0{,}8051 - (1 - 0{,}8051) = 0{,}6102$ R

<u>Svar</u>: Slh att nettovikten ligger mellan 502 och 504 gram är ca 0,6102

4

c) X = nettovikten för en låda bestående av 10 paket flingor



Y = nettovikt för ett flingpaket

$X = \sum_{i=1}^{10} Y_i$ medelnettovikten $= \frac{X}{10} = \frac{\sum_{i=1}^{10} Y_i}{10}$

$E(X) = E\left(\sum_{i=1}^{10} Y_i\right) = \sum_{i=1}^{10} E(Y_i) = 10 \cdot E(Y_i) =$

$= 10 \cdot 503 = 5030 \Rightarrow E\left(\frac{X}{10}\right) = \frac{5030}{10} = 503$ R

$\sigma(X) = \sigma\left(\sum_{i=1}^{10} Y_i\right) = \sqrt{\sum_{i=1}^{10} \sigma^2(Y_i)} = \sqrt{10} \cdot \frac{50}{43} = \frac{500}{43}$

(ingen kovarians pga oberoende mellan alla Y_i) Bra

$\Rightarrow \sigma\left(\frac{X}{10}\right) = \frac{500}{43} / 10 = \frac{50}{43}$

<u>Svar</u>: Väntevärdet & standardavvikelsen för medelnettovikten är 503 respektive 50/43

SU, STATISTIK

Skrivsal: Ugglevik Anonymkod: ST-0001 Blad nr: 4

③ c) forts.

Vidare förklaring av c):

Vi vet att alla Y_i är normal fördelade vilket gör att $\sum_{i=1}^{10} y_i = X$ också kommer vara normal fördelad (pga linjär kombination av Y_i). För X gäller det dessutom att $\mu = \mu_{Y_1} + \dots \mu_{Y_{10}} = 5030$ och $\sigma = \sigma_{Y_1} + \dots + \sigma_{Y_{10}} = \frac{500}{43}$. På samma sätt kan $\frac{X}{10}$ tänkas ha $\mu = \frac{\mu_{Y_1}}{10} + \dots \frac{\mu_{Y_{10}}}{10} = 503$ och $\sigma = \frac{\sigma_{Y_1}}{10} + \dots + \frac{\sigma_{Y_{10}}}{10} = \frac{50}{43}$ R 6

detta ger för d):

$$P\left(502 < \frac{X}{10} < 504\right) = P\left(\frac{502-503}{50/43} < Z < \frac{504-503}{50/43}\right)$$

vilket vi i uppgift b) såg blev 0,6102 (se b) för uträkning).

__Svar:__ slh att medelnettovikten för paketen i en låda är mellan 502 och 504 gram är ca 0,6102

Följdfel 2

(17)

(4) Y = hastigheten hos en gasmolekyl

$$f(y) = \begin{cases} 2ye^{-y^2}, & y \geq 0 \\ 0, & \text{annars} \end{cases}$$

U = kinetiska energin hos en gasmolekyl

$$U = 6Y^2$$

a) (1) Definitionsområdet för u:

$Y=0 \Rightarrow u=0$ $\quad y=1 \Rightarrow u=6$ $\quad y=2 \Rightarrow u=24$

Då Y bara är definierat för positiva värden är U en strängt växande funktion av Y med definitionsmängd: R

$$0 \leq u < \infty$$

(2) Inversfunktionen $h^{-1}(u)$:

Då U är en strängt växande funktion av Y kan transformationsmetoden användas. R

Vi tar därför fram inversfunktionen av U, dvs $Y = \left(\frac{u}{6}\right)^{1/2} = \frac{u^{1/2}}{6^{1/2}} = \frac{1}{6^{1/2}} \cdot u^{1/2} = h^{-1}(u)$ R

(3) Derivatan av $h^{-1}(u)$:

$$\frac{dh^{-1}(u)}{du}\left(\frac{1}{6^{1/2}} \cdot u^{1/2}\right) = \frac{1}{2} \cdot \frac{1}{6^{1/2}} \cdot u^{1/2-1} = \frac{1}{2\sqrt{6}} \cdot u^{-1/2} =$$

$$= \frac{1}{2\sqrt{6}} \cdot \frac{1}{\sqrt{u}} = \frac{1}{2\sqrt{6}\sqrt{u}} = \frac{1}{2 \cdot 6^{1/2} \cdot u^{1/2}}$$ R

SU, STATISTIK

Skrivsal: Ugglevik **Anonymkod:** ST-0001 **Blad nr:** 5

4 a, forts.

$$f_U(u) = f_Y\left[h^{-1}(u)\right]\left|\frac{du^{-1}}{du}\right| =$$

$$= 2\left(\frac{u^{1/2}}{6^{1/2}}\right)e^{-\left(\frac{u^{1/2}}{6^{1/2}}\right)^2} \cdot \frac{1}{2\cdot 6^{1/2}\cdot u^{1/2}} =$$

$$= \frac{\dfrac{2u^{1/2}e^{-\left(\frac{u}{6}\right)}}{6^{1/2}}}{\dfrac{2\cdot 6^{1/2}\cdot u^{1/2}}{1}} = \frac{2u^{1/2}e^{-\left(\frac{u}{6}\right)}}{6^{1/2}} \cdot \frac{1}{2\cdot 6^{1/2}\cdot u^{1/2}} =$$

$$= \frac{2u^{1/2}e^{-\left(\frac{u}{6}\right)}}{6^{1/2}\cdot 2\cdot 6^{1/2}\cdot u^{1/2}} = \frac{\cancel{2}\cdot \cancel{u^{1/2}}\cdot e^{-\left(\frac{u}{6}\right)}}{\cancel{2}\cdot \cancel{u^{1/2}}\cdot 6} = \frac{e^{-\left(\frac{u}{6}\right)}}{6} = \underline{\underline{\frac{1}{6}\cdot e^{-u/6}}} \quad R$$

Så: $f_U(u) = \begin{cases} \frac{1}{6}e^{-u/6} , & 0 \leq u \\ 0 , & \text{annars} \end{cases}$ TABELL 1 R

$$F_U(u) = \int_0^u \frac{1}{6}e^{-u/6}\,du = \frac{1}{6}e^{-\frac{1}{6}\cdot u}\,du = \left[\frac{1}{-1/6}\cdot\frac{1}{6}e^{-\frac{1}{6}u}\right]^u =$$

$$= \left[-e^{-u/6}\right]_0^u = -e^{-u/6} - \left(-e^{-0/6}\right) = -e^{-u/6} + e^0 = \underline{\underline{1 - e^{-u/6}}} \quad R$$

Så: $F_U(u) = \begin{cases} 0, & u < 0 \\ 1 - e^{-u/6}, & u \geq 0 \end{cases}$ TABELL 2

R

<u>Svar</u>: Se Tabell 1 för täthetsfunktionen och Tabell 2 för fördelnings-funktionen för den kinetiska energin (U).

18

b)

<u>Svar</u>: Den kinetiska energin är exponentialfördelad med $\beta = 6$

R

(Den generella tätheten $\frac{1}{\beta} e^{-y/\beta}$ är i vårt fall $\frac{1}{6} e^{-u/6}$)

2

20

5

Y_1 = tid från att kunden ställer sig i kö tills att ärendet är klart

Y_2 = kötiden

$$f(y_1, y_2) = \begin{cases} e^{-y_1}, & 0 < y_2 < y_1 < \infty \\ 0, & \text{annars} \end{cases}$$

a) $$f_1(y_1) = \int_0^{y_1} e^{-y_1} \, dy_2 = \left[y_2 e^{-y_1} \right]_0^{y_1} = \begin{cases} y_1 e^{-y_1}, & 0 < y_1 < \infty \\ 0, & \text{annars} \end{cases}$$

R

TABELL 1

SU, STATISTIK

Skrivsal: Ugglevik Anonymkod: ST-0001 Blad nr: 6

5 a, forts.

$$f_2(y_2) = \int_{y_2}^{\infty} e^{-y_1}\, dy_1 = \left[-e^{-y_1}\right]_{y_2}^{\infty} = -e^{-\infty} - (-e^{-y_2}) =$$

$$= 0 + e^{-y_2}$$ Så: $f_2(y_2) = \begin{cases} e^{-y_2}, & 0 < y_2 < \infty \\ 0, & \text{annars} \end{cases}$ TABELL 2

R

Svar: För marginalfördelningarna för Y_1 & Y_2, se Tabell 1, respektive Tabell 2

6

b, Nej, Y_1 & Y_2 är inte stokastiskt oberoende. Detta går att se direkt i definitionsområdet där $Y_2 < Y_1$ eller på att produkten av marginalfördelningarna inte är lika med den simultana täthetsfunktionen. Intuitivt går det att tänka att den totala tiden är beroende av kötiden.

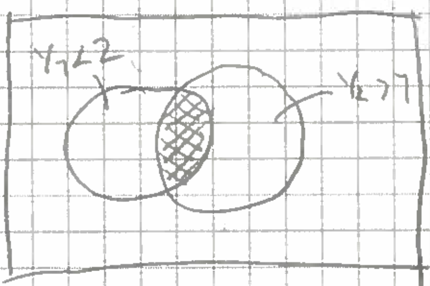
R

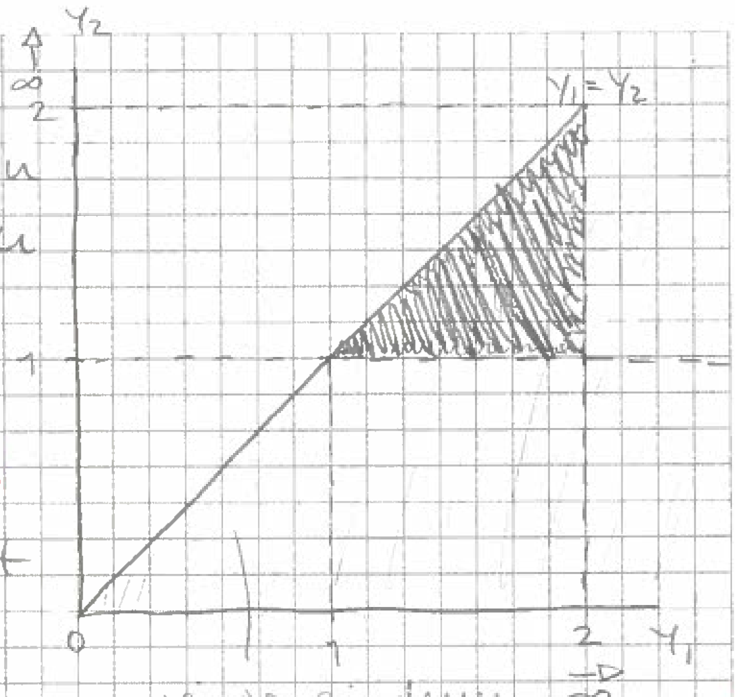
R

2

c) Att den totala tiden understiger 2 min samtidigt som kötiden överstiger 1 min borde motsvara $P(Y_1<2 \cap Y_2>1)$ R

då man kan se det som:





Vi är på denna halva $Y_1=Y_2$ då det är givet i frågan att $Y_1>Y_2$ R

och att det är snittet, dvs när båda händelserna $Y_1<2$ & $Y_2>1$ inträffar vi är intresserade av. Om vi enligt diagrammet ovan ser det området ovanifrån kan vi se att Y_1 går från linjen ($Y_1=Y_2$) till 2 och att Y_2 går från 1 till 2 (om Y_1 är inre integral, vilket vi vill för att slippa partiell integration). Detta ger oss: Bra!

$$\int_1^2\int_{y_2}^2 e^{-y_1}\,dy_1\,dy_2 = \int_1^2 \left[-e^{-y_1}\right]_{y_2}^2 dy_2 = \int_1^2 \left(-e^{-2}-(-e^{-y_2})\right) = \int_1^2 -e^{-2}+e^{-y_2}\,dy_2$$ R

$$= \left[-y_2e^{-2}+(-e^{-y_2})\right]_1^2 = \left(-2e^{-2}-e^{-2}\right)-\left(-e^{-2}-e^{-1}\right) =$$

$$= -2e^{-2}-\cancel{e^{-2}}+\cancel{e^{-2}}+e^{-1} = e^{-1}-2e^{-2} \approx 0{,}0972$$ R

<u>Svar</u>: Sln att $Y_1<2$ samtidigt som $Y_2>1$ är $\approx 0{,}0972$ 12

20

Genomgående tydliga och väl motiverade lösn., bra!