



Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF1910/SF1925 TILLÄMPAD STATISTIK,
TISDAG 22 APRIL 2025 KL 8.00–13.00.

Examinator: Björn-Olof Skytt, 08-790 86 49

Tillåtna hjälpmedel: Formel- och tabellsamling i Matematisk statistik (utdelas vid tentamen), miniräknare.

Tentamen består av två delar, benämnda del I och del II.

Del I består av uppgifterna 1–12 och varje korrekt svar ger 1 poäng. På denna del ska endast svar anges, i form av val av ett av de möjliga svarsalternativen. Svaren ska anges på svarsblanketten (utdelas vid tentamen). Studenter som är godkända på kontrollskrivningen behöver ej besvara uppgift 1–3, utan får tillgodoräkna sig dessa tre uppgifter (i svarsblanketten anges ordet "Bonus"). Studenter som är godkända på den andra datorlaborationen behöver ej besvara uppgift 12, utan får tillgodoräkna sig denna uppgift (i svarsblanketten anges ordet "Bonus"). Gränsen för godkänt är 9 poäng. Möjlighet att komplettera ges för tentander med 8 poäng.

Del II består av uppgifterna 13–16 och varje korrekt lösning ger 10 poäng. Del II rättas bara för studenter som är godkända på eller får komplettera del I och poäng på del II krävs för högre betyg än E. På denna del ska resonemang och uträkningar vara så utförliga och väl motiverade att de är lätta att följa. Införda beteckningar ska förklaras och definieras samt numeriska svar ska anges med minst tre värdesiffrors noggrannhet. Studenter som är godkända på den andra datorlaborationen får 3 bonuspoäng på del II.

Tentamen kommer att vara rättad inom tre arbetsveckor från skrivningstillfället och kommer att finnas tillgänglig på studentexpeditionen minst sju veckor efter skrivningstillfället.

Del I

Uppgift 1

Låt $P(A) = 0.30$, $P(A \mid B) = 0.25$ och $P(A^* \cap B^*) = 0.40$. Beräkna $P(B)$.

A: 0.4

B: 0.6

C: 0.3

D: 0.375

Uppgift 2

Låt X och Y vara oberoende stokastiska variabler med fördelningar $\text{Bin}(10, 0.9)$ respektive $\text{Bin}(5, 0.9)$. Beräkna $P(X = 12 - Y)$.

A: 0.028

B: 0.064

C: 0.129

D: 0.003

Uppgift 3

Låt X och Y vara stokastiska variabler så att

$$E(X) = 2 \quad E(Y) = 3 \quad C(X, Y) = 1 \quad D(X) = 3 \quad D(Y) = 2.$$

Beräkna $E((2X + 1)(3Y + 2))$.

A: 25

B: 42

C: 55

D: 61

Uppgift 4

För att underlätta återfinnandet av borttappade nycklar kan man beställa en nyckelbricka med ett unikt id-nummer som sedan kan spåras tillbaka till ägaren. Enligt marknadsföringen hos ett säkerhetsföretag så är 9 av 10 återfunna nycklar märkta med deras nyckelbricka. Antag att 1 av 20 borttappade nycklar återfinns (oavsett märkning) och att 1 av 7 borttappade nycklar är märkta. Bestäm kvoten mellan sannolikheten att en märkt nyckel återfinns jämfört med en omärkt nyckel (enligt marknadsföringen).

A: 19

B: 54

C: 9

D: 1

Uppgift 5

Låt X och Y vara oberoende stokastiska variabler med fördelningarna $N(0.8, 1.2)$ respektive $N(0.7, 1.1)$. Vad är $P(3X + 2Y < 3.3)$?

A: 0.47

B: 0.49

C: 0.55

D: 0.45

Uppgift 6

Antag att $X_1, X_2, \dots, X_{96}$ är oberoende och likafördelade med täthetsfunktion

$$f(x) = \begin{cases} 3x^2, & x \in [0, 1] \\ 0, & \text{annars.} \end{cases}$$

Beräkna den approximativa sannolikheten att

$$\sum_{i=1}^{96} X_i < 75.$$

A: 0.56

B: 0.48

C: 0.94

D: 0.98

Uppgift 7

Två undersökningar gjordes på två grupper av patienter som hade blivit vaccinerade respektive inte hade blivit vaccinerade. Fem hundra vaccinerade undersöktes där fyrtionio hade fått vinterkräksjuka. I den icke vaccinerade gruppen undersöktes sex hundra där femtioåtta hade fått vinterkräksjuka. Låt p_1 stå för andelen patienter som fått vinterkräksjuka trots att de har vaccinerats och låt p_2 stå för andelen patienter som fått vinterkräksjuka utan att ha vaccinerats. Vi är intresserade av parametern $p_1 - p_2$. Bestäm medelfelet för skattningen av $p_1 - p_2$.

A: 0.018

B: 0.419

C: 0.297

D: 0.441

Uppgift 8

En stokastisk variabel X har täthetsfunktionen

$$f_X(x) = \begin{cases} \frac{\theta}{(1+x)^{\theta+1}} & \text{för } x \geq 0 \\ 0 & \text{för } x < 0, \end{cases}$$

där $\theta > 0$ är en okänd parameter som antingen har värdet 1, 2, 3, eller 4.

Man har ett slumpmässigt stickprov från denna fördelning omfattande tre värden: 3, 5, och 9. Ange maximum-likelihood-skattningen av θ .

A: 1

B: 2

C: 3

D: 4

Uppgift 9

För 15 år sedan ville man på ett lärosäte i ett annat land undersöka om studenterna ägnade mer tid åt att surfa på internet som att titta på TV. Man gjorde en undersökning och fick resultat enligt tabellen nedan.

Person nr.	1	2	3	4	5	6	7	8
Surftid per vecka (h)	2	7	3	8	9	15	7	2
TV-tid per vecka (h)	4	15	5	3	4	4	4	8

Vi kan här anta att observationerna är normalfördelade. Bilda ett ensidigt nedåt begränsat konfidensintervall med en konfidensgrad på 95% för hur många fler timmar studenter spenderar på internetsurfande jämfört med TV-tittande.

A: $(-3.0, \infty)$

B: $(-3.5, \infty)$

C: $(-3.7, \infty)$

D: $(-4.6, \infty)$

Uppgift 10

På ett lärosäte i ett annat land ville man undersöka om det var någon skillnad mellan män och kvinnor vad gällde hur lång tid det tog för dem att genomföra läkarutbildningen vid lärosätet. Man utförde två oberoende stickprov och fick resultat enligt tabellen nedan.

	Män	Kvinnor
Stickprovsstorlek	100	85
Medelvärde (år)	8.8	8.4
Stickprovsstandardavvikelse (år)	1.4	1.5

Observationerna kan här inte anses vara utfall av normalfördelade stokastiska variabler, men stickproven kan anses vara stora. Bilda ett dubbelsidigt konfidensintervall med en approximativ konfidensgrad på 95% för skillnaden i studietid mellan män och kvinnor och ange dess övre gräns.

A: 0.57

B: 0.61

C: 0.75

D: 0.82

Uppgift 11

I en marknadsundersökning utfrågades ett urval av 120 konsumenter om vilken typ av förpackning de föredrog för en ny produkt. De hade fyra olika typer att välja mellan. 39 föredrog typ A, 27 föredrog typ B, 24 föredrog typ C och resten föredrog typ D. Kan vi förkasta nollhypotesen H_0 att alla fyra typer av förpackningar är lika populära?

A: H_0 kan varken förkastas på risknivån 1% eller risknivån 5%B: H_0 kan både förkastas på risknivån 1% och risknivån 5%C: H_0 kan förkastas på risknivån 1%, men inte på risknivån 5%D: H_0 kan förkastas på risknivån 5%, men inte på risknivån 1%**Uppgift 12**

Vid enkel linjär regression gäller följande samband

$$y_i = \alpha + \beta x_i + \varepsilon_i,$$

där man observerat paren (x_i, y_i) , $i = 1, \dots, n$ och ε_i betecknar de slumpmässiga felen. Man önskar testa nollhypotesen $H_0 : \beta = 0$ mot $H_1 : \beta \neq 0$ och får P -värdet till 0.042. Vilket av följande påståenden är sant?

A: Vi förkastar H_0 och kan därmed inte förkasta att y_i :na beror av x_i :na på risknivån 1%B: Vi förkastar H_0 och kan därmed inte förkasta att y_i :na beror av x_i :na på risknivån 5%C: Vi förkastar H_0 och kan därmed förkasta att y_i :na beror av x_i :na på risknivån 1%D: Vi förkastar H_0 och kan därmed förkasta att y_i :na beror av x_i :na på risknivån 5%

Del II

Uppgift 13

Busshållplatsen vid korsningen Valhallavägen–Odengatan i Stockholm har plats för att två bussar ska kunna stanna samtidigt. Vid denna busshållplats stannar fem olika bussar: två stombusslinjer (linje 4 och 6, även kallade ”blåbussar”) samt tre vanliga busslinjer (linje 61, 67 och 72). Antag att bussarna ankommer till hållplatsen slumpmässigt och oberoende av varandra. Under rusningstid (kl. 07–09) har stombusslinjerna en turtäthet på 8 bussar i timmen, medan linjerna 61, 67 och 72 har turtätheter på 5, 3 respektive 3 bussar per timme. Antag vidare att varje buss stannar i en minut vid busshållplatsen och att en buss precis har stannat. Vad är sannolikheten att det blir för trångt på busshållplatsen? (10 p)

Uppgift 14

Låt X vara exponentialfördelad med parameter λ och låt $Y = \lceil X \rceil$ vara X avrundat uppåt. Med andra ord har vi

$$Y = \begin{cases} 1, & \text{om } 0 \leq X \leq 1, \\ 2, & \text{om } 1 < X \leq 2, \\ 3, & \text{om } 2 < X \leq 3, \\ \text{osv.} \end{cases}$$

Bestäm Y :s sannolikhetsfunktion. Kan den beskrivas som en av fördelningarna i formelsamlingen? Om så är fallet, vad heter fördelningen? (10 p)

Uppgift 15

Den naturliga bakgrundsstrålningen (uttryckt som antalet registrerade pulser per sekund i Becquerel, förkortat Bq) har en intensitet av $\lambda = 1 \text{ s}^{-1}$ och antalet registrerade pulser under ett tidsintervall (t_1, t_2) beskrivs väl av en Poissonfördelning med väntevärde $\lambda(t_2 - t_1)$. På grund av en kärnkraftsolycka i ett avlägset land misstänker man att intensiteten här hos oss har ökat.

- Antag att man mäter under 15 sekunder och därvid registrerar 20 pulser. Undersök med lämpligt test eller konfidensintervall om man på risknivå 5% kan förkasta att intensiteten inte överstiger den naturliga bakgrundsstrålningen. (4 p)
- Antag att intensiteten i själva verket har ökat till $\lambda = 1.2 \text{ s}^{-1}$. Hur länge måste man då mäta för att sannolikheten ska vara 50% att upptäcka detta om man gör ett test på en risknivå på approximativt 5%. (6 p)

Uppgift 16

Antag att Z_1, Z_2, Z_3, Z_4 är oberoende stokastiska variabler som alla är likformigt fördelade enligt $U(a, b)$. Sätt $X = \min(Z_1, Z_2)$ och $Y = \max(Z_3, Z_4)$. Anta att X får utfallet x och att Y får utfallet y . Härled och ta fram skattningarna av parametrarna a och b med hjälp av minstakvadrat-metoden. Skattningarna ska anges som funktioner av x och y . (10 p)

Lycka till!



Avd. Matematisk statistik

KTH Matematik

LÖSNINGSFÖRSLAG TENTAMEN I SF1910/SF1925
TILLÄMPAD STATISTIK,
TISDAG 22 APRIL 2025 KL 8.00–13.00.

Del I, Svar.

1. A
2. C
3. D
4. B
5. D
6. C
7. A
8. A
9. B
10. D
11. A
12. B

Del I, Lösningsförslag.**Uppgift 1**

Först har vi

$$P(A \cap B) = P(A | B) P(B) = 0.25P(B)$$

och sedan noterar vi att

$$P(A \cup B) = 1 - P(A^* \cap B^*) = 1 - 0.4 = 0.6,$$

samt att

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - 0.25P(B) = P(A) + 0.75P(B).$$

Stoppar vi in värdena på $P(A \cup B)$ samt $P(A)$ får vi då

$$P(B) = \frac{0.6 - 0.3}{0.75} = 0.4.$$

Uppgift 2

Eftersom

$$P(X = 12 - Y) = P(X + Y = 12)$$

samt att $X + Y$ har fördelningen $\text{Bin}(10 + 5, 0.9) = \text{Bin}(15, 0.9)$ får vi att

$$P(X = 12 - Y) = \binom{15}{12} 0.9^{12} 0.1^3 \approx 0.129.$$

Uppgift 3

$$\begin{aligned} E((2X + 1)(3Y + 2)) &= E(6XY + 4X + 3Y + 2) = 6E(XY) + 4E(X) + 3E(Y) + 2 \\ &= 6(C(X, Y) + E(X)E(Y)) + 4E(X) + 3E(Y) + 2 \\ &= 6(1 + 2 \cdot 3) + 4 \cdot 2 + 3 \cdot 3 + 2 = 61 \end{aligned}$$

Uppgift 4

Låt A vara händelsen att en borttappad nyckel är märkt medan B är händelsen att den återfinns. Enligt uppgiftslydelsen har vi då att $P(A | B) = 9/10$ (från marknadsföringen) och $P(A) = 1/7$. Vi söker nu $P(B | A) / P(B | A^*)$, vilket kan beräknas genom Bayes sats till

$$\frac{P(B | A)}{P(B | A^*)} = \frac{\frac{P(A|B)P(B)}{P(A)}}{\frac{P(A^*|B)P(B)}{P(A^*)}} = \frac{P(A^*)}{P(A)} \cdot \frac{P(A | B)}{P(A^* | B)} = \frac{6/7}{1/7} \cdot \frac{9/10}{1/10} = 6 \cdot 9 = 54.$$

Uppgift 5

Summan $3X + 2Y$ är normalfördelad med väntevärde $3E(X) + 2E(Y) = 3 \cdot 0.8 + 2 \cdot 0.7 = 3.8$ samt standardavvikelse $\sqrt{3^2 V(X) + 2^2 V(Y)} = \sqrt{3^2 \cdot 1.2^2 + 2^2 \cdot 1.1^2} = \sqrt{15.28}$. Vi får då

$$\begin{aligned} P(3X + 2Y < 3.3) &= P\left(\frac{3X + 2Y - 3.8}{\sqrt{15.28}} < \frac{3.3 - 3.8}{\sqrt{15.28}}\right) \\ &= \Phi\left(-\frac{0.5}{\sqrt{15.28}}\right) \approx \Phi(-0.12) = 1 - \Phi(0.12) \approx 1 - 0.5478 = 0.4522. \end{aligned}$$

Uppgift 6

Vi ser först att

$$E(X_i) = \int_0^1 x \cdot 3x^2 dx = \frac{3}{4}$$

samt att

$$E(X_i^2) = \int_0^1 x^2 \cdot 3x^2 dx = \frac{3}{5},$$

vilket ger

$$V(X_i) = E(X_i^2) - E(X_i)^2 = \frac{3}{80}$$

så vi får $D(X_i) = \sqrt{3/80}$.

Med hjälp av centrala gränsvärdessatsen får vi att summan är approximativt normalfördelad med väntevärde $96 \cdot 3/4 = 72$ samt standardavvikelse $\sqrt{96 \cdot 3/80} = \sqrt{3.6}$. Vi får då sannolikheten

$$P\left(\sum_{i=1}^{96} X_i < 75\right) \approx \Phi\left(\frac{75 - 72}{\sqrt{3.6}}\right) \approx \Phi(1.58) \approx 0.9429.$$

Uppgift 7

$$\sqrt{\frac{\frac{x_1}{n_1} \left(1 - \frac{x_1}{n_1}\right)}{n_1} + \frac{\frac{x_2}{n_2} \left(1 - \frac{x_2}{n_2}\right)}{n_2}} = \sqrt{\frac{\frac{49}{500} \left(1 - \frac{49}{500}\right)}{500} + \frac{\frac{58}{600} \left(1 - \frac{58}{600}\right)}{600}} = 0.01795352 = 0.0180$$

Uppgift 8

Se §9.1 i formelsamlingen. Vi ska hitta det som maximerar $L(\theta)$ där θ kan anta värdena 1, 2, 3 eller 4.

$$\begin{aligned} L(\theta) &= f_{X_1}(x_1) f_{X_2}(x_2) f_{X_3}(x_3) = \frac{\theta}{(1+x_1)^{\theta+1}} \frac{\theta}{(1+x_2)^{\theta+1}} \frac{\theta}{(1+x_3)^{\theta+1}} = \\ &= \frac{\theta}{(1+3)^{\theta+1}} \frac{\theta}{(1+5)^{\theta+1}} \frac{\theta}{(1+9)^{\theta+1}} = \frac{\theta^3}{20^{\theta+1}} \end{aligned}$$

$L(1) = \frac{1}{20^2}$, $L(2) = \frac{8}{20^3}$, $L(3) = \frac{27}{20^4}$, $L(4) = \frac{64}{20^5}$. $L(1)$ är störst av dessa fyra värden. Alltså är $\theta = 1$ maximum-likelihood-skattningen av θ .

Uppgift 9

Här har vi parvisa observationer, dvs. stickprov i par. Vi bildar då ett konfidensintervall med hjälp av §12.2 och §11.1(d) i formelsamlingen och får efter att vi bytt ut $\frac{\alpha}{2}$ mot α i nedre gränsen

$$I_{\mu_z} = \bar{z} - t_{0.05}(7) \cdot \frac{s_z}{\sqrt{n}} = 0.75 - 1.89 \cdot \frac{6.36}{\sqrt{8}} \approx -3.5$$

Uppgift 10

Eftersom stickproven kan anses vara stora så kan vi på grund av centrala gränsvärdessatsen anse att $\bar{X}$ och $\bar{Y}$ är approximativt normalfördelade och därmed kan vi med hjälp av §12.3 i F.S. bilda ett konfidensintervall med approximativ konfidensgrad eftersom $\theta^* = \bar{X} - \bar{Y}$ i §12.3 då är approximativt normalfördelad.

Då får vi

$$I_{\mu_x - \mu_y} = \bar{x} - \bar{y} \pm \sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}} \lambda_{\frac{\alpha}{2}} = 8.8 - 8.4 \pm \sqrt{\frac{1.4^2}{100} + \frac{1.5^2}{85}} 1.96 = 0.4 \pm 0.42$$

Alltså är den övre gränsen 0.82.

Uppgift 11

Här har vi test av given fördelning och använder oss av § 14.3 i formelsamlingen, vilket ger

$$Q = \sum_{i=1}^k \frac{(x_i - np_i)^2}{np_i}$$

Här är $n = 120$, $k = 4$ och alla $p_i = 0.25$, så alla $np_i \geq 5$. Detta ger då

$$Q = \frac{(39 - 30)^2 + (27 - 30)^2 + (24 - 30)^2 + (30 - 30)^2}{30} = 4.2$$

Till slut får vi från tabellsamlingen att $\chi_{0.05}^2(3) = 7.81$ och $\chi_{0.01}^2(3) = 11.3$. Eftersom $Q = 4.2$ är mindre än båda dessa värden kan H_0 varken förkastas på risknivån 5% eller 1%.

Uppgift 12

Eftersom p-värdet = 0.042 är mindre än 5% men större än 1% kan vi förkasta H_0 på risknivån 5%, men inte på risknivån 1%. Om vi förkastar att $\beta = 0$ kan vi inte förkasta att y_i :na beror av x_i :na på risknivån 5%. Alltså är B rätt svarsalternativ.

Del II, Lösningsförslag:**Uppgift 13**

Eftersom bussarna anländer slumpmässigt och oberoende av varandra följer de en Poissonmodell. Låt X_i vara antalet bussar på linje i som anländer under en tidsperiod av längd t (mätt i timmar). Vi har då att X_4 och X_6 har fördelningen $\text{Po}(8t)$, X_{61} har fördelningen $\text{Po}(5t)$ samt att X_{67} och X_{72} har fördelningen $\text{Po}(3t)$. Alla dessa X_i är oberoende av varandra (bussarna anländer oberoende av varandra, som det står i beskrivningen). Antalet bussar totalt som stannar under tidsperioden är då $Y = X_4 + X_6 + X_{61} + X_{67} + X_{72}$, vilket har fördelningen $\text{Po}(8t + 8t + 5t + 3t + 3t) = \text{Po}(27t)$. Om en buss precis har stannat får det plats en buss till innan det blir för trångt. Med andra ord blir det för trångt om två eller fler bussar stannar under tiden den första bussen står där, vilket är en minut ($1/60$ timme). Vi sätter då $t = 1/60$ och den sökta sannolikheten blir då

$$\begin{aligned} P(Y \geq 2) &= 1 - P(Y = 0) - P(Y = 1) = 1 - \left(\frac{(27/60)^0}{0!} + \frac{(27/60)^1}{1!} \right) e^{-27/60} \\ &= 1 - (1 + 27/60)e^{-27/60} \approx 0.0754. \end{aligned}$$

Uppgift 14

Vi har alltså täthetsfunktionen

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & \text{annars.} \end{cases}$$

Sannolikhetsfunktionen för Y ges av

$$\begin{aligned} p_Y(y) &= P(Y = y) = P([X] = y) = P(y - 1 < X \leq y) \\ &= \int_{y-1}^y f_X(x) dx \\ &= \int_{y-1}^y \lambda e^{-\lambda x} dx \\ &= [-e^{-\lambda x}]_{y-1}^y \\ &= e^{-\lambda(y-1)} - e^{-\lambda y} = e^{-\lambda(y-1)}(1 - e^{-\lambda}). \end{aligned}$$

Om vi sätter $p = 1 - e^{-\lambda}$ kan vi skriva detta på formen

$$p_Y(y) = p(1 - p)^{y-1}.$$

Med andra ord är Y $\text{ffg}(1 - e^{-\lambda})$ -fördelad.

Uppgift 15

Låt $X(t)$ vara antalet registrerade partiklar under $(0, t)$. Med andra ord är $X(t)$ Po(λt)-fördelad.

(a) Här har vi $H_0 : \lambda = 1$ och $H_1 : \lambda \geq 1$.

Om H_0 är sann och $t = 15$ s så gäller att $X(15) \in \text{Po}(15)$. Eftersom vi observerat $x = 20$ blir p-värdet $= P(X \geq 20) = 1 - P(X \leq 19) = [\text{se tabell 5}] = 1 - 0.87522 = 0.1248$. Eftersom $0.1248 > \alpha = 0.05$ så kan inte H_0 förkastas på risknivån 5%.

(b) $X(t)$ är approximativt $N(\lambda t, \sqrt{\lambda t})$ vilket under H_0 blir $N(t, \sqrt{t})$. Styrkan hos testet är sannolikheten att förkasta $H_0 : \lambda = 1$ om $H_1 : \lambda = 1.2$ är sann. Vi har alltså här att

$$\begin{aligned} h(1.2) &= 0.5 = P(X(t) > t + \sqrt{t}\lambda_{0.05} \mid \lambda = 1.2) \\ &= P\left(\frac{X(t) - 1.2t}{\sqrt{1.2t}} > \frac{t + \sqrt{t}\lambda_{0.05} - 1.2t}{\sqrt{1.2t}}\right) \\ &= 1 - \Phi\left(\frac{\sqrt{t}\lambda_{0.05} - 0.2t}{\sqrt{1.2t}}\right) = 0.5. \end{aligned}$$

Dvs. $\sqrt{t}\lambda_{0.05} = 0.2t \Rightarrow \lambda_{0.05} = 0.2\sqrt{t} \Rightarrow t = (1.6449 \cdot 5)^2 = 67.64$ s.

Uppgift 16

Eftersom $Y = \max(Z_3, Z_4)$ får vi dess fördelningsfunktion till

$$F_Y(y) = \left(\frac{y-a}{b-a}\right)^2,$$

vilket ger täthetsfunktionen

$$f_Y(y) = 2 \left(\frac{y-a}{b-a}\right) \frac{1}{b-a}.$$

Vi vill nu integrera

$$yf_Y(y) = (y-a) \cdot 2 \left(\frac{y-a}{b-a}\right) \frac{1}{b-a} + a \cdot 2 \left(\frac{y-a}{b-a}\right) \frac{1}{b-a}$$

över intervallet $[a, b]$ för att få $E(Y)$. För att förenkla beräkningarna gör vi variabelbytet $z = y-a$, vilket ger

$$E(Y) = \int_0^{b-a} \frac{2}{(b-a)^2} z^2 + \frac{2a}{(b-a)^2} z dz = \frac{2(b-a)}{3} + a = \frac{a}{3} + \frac{2b}{3}$$

På samma sätt har vi att $X = \min(Z_1, Z_2)$ ger

$$F_X(x) = 1 - \left(1 - \left(\frac{x-a}{b-a}\right)\right) \left(1 - \left(\frac{x-a}{b-a}\right)\right) = 2 \left(\frac{x-a}{b-a}\right) - \left(\frac{x-a}{b-a}\right)^2$$

samt

$$f_X(x) = \frac{2}{(b-a)} - \frac{2(x-a)}{(b-a)^2}.$$

Som innan vill vi integrera

$$xf_X(x) = \frac{2x}{(b-a)} - \frac{2(x-a)^2}{(b-a)^2} - \frac{2a(x-a)}{(b-a)^2}$$

över intervallet $[a, b]$ för att få $E(X)$. Med variabelbytet $z = x - a$ får vi

$$E(X) = \int_0^{b-a} \frac{2(z+a)}{(b-a)} - \frac{2z^2}{(b-a)^2} - \frac{2az}{(b-a)^2} dz = b + a - \frac{2}{3}(b-a) - a = \frac{b}{3} + \frac{2a}{3}$$

Vi kan nu ställa upp kvadratsumman och får

$$Q = \left(y - \left(\frac{a}{3} + \frac{2b}{3} \right) \right)^2 + \left(x - \left(\frac{b}{3} + \frac{2a}{3} \right) \right)^2$$

Deriverar vi med avseende på a och b och sätter till noll får vi den kritiska punkten

$$(a, b) = (2x - y, 2y - x)$$

. Eftersom Q är kvadratisk i a och b ser vi att punkten utgör ett global minimum, vilket ger

$$a_{\text{obs,MK}}^* = 2x - y \text{ och } b_{\text{obs,MK}}^* = 2y - x.$$