



Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF1923 OCH SF1924 SANNOLIKHETSTEORI OCH STATISTIK,
SAMT TENTAMEN I SF1935 SANNOLIKHETSTEORI OCH STATISTIK MED TILLÄMPNING
INOM MASKININLÄRNING
MÅNDAG 26 MAJ 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 då ordet ”Bonus”). Studenter som är godkända på den andra datorlaborationen behöver ej besvara uppgift 12, utan får tillgodoräkna sig denna uppgift (i svarsblanketten anges då ordet ”Bonus”). Dessa tillgodoräknanden gäller för den här tentamen och vid omtentamen i augusti 2025. 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 fyra* värdesiffrors noggrannhet. Studenter som är godkända på den andra datorlaborationen får dessutom tre bonuspoäng på del II. Dessa bonuspoäng gäller för den här tentamen och vid omtentamen i augusti 2025.

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

För händelserna A och B gäller att ett av de fyra påståendena nedan är felaktigt. Vilket?

A: $P(B|A) = 1 - P(B^*|A)$

B: $P(A \cap B) = P(B) \cdot P(A|B)$

C: $P(A^* \cup B^*) = 1 - P(A) - P(B) + P(A \cap B)$

D: $P(B \cap A^*) = 1 - P(A) - P(A^* \cap B^*)$

Uppgift 2

En stokastisk variabel (X, Y) har täthetsfunktionen

$$f_{X,Y}(x, y) = \begin{cases} cx, & x \geq 0, y \geq 0, x + y \leq 1, \\ 0, & \text{annars.} \end{cases}$$

Bestäm konstanten c .

A: 6

B: 5

C: 3

D: 2

Uppgift 3

Låt X och Y vara stokastiska variabler sådana att $V(X) = 2$, $V(Y) = 2$, och $V(X + 3Y) = 8$. Beräkna korrelationskoefficienten $\rho(X, Y)$.

A: 0

B: $-\frac{1}{2}$

C: -1

D: -2

Uppgift 4

Den stokastiska variabeln ξ tillhör Exponentialfördelningen med väntevärdet 200 timmar. Bestäm fördelningens median.

A: 921

B: 100

C: 139

D: 184

Uppgift 5

Låt X_1 och X_2 vara oberoende stokastiska variabler sådana att $X_1 \in N(2, 1)$ och $X_2 \in N(0, 2)$, d.v.s. $E(X_1) = 2$, $V(X_1) = 1$, $E(X_2) = 0$, $V(X_2) = 4$. Låt $Y = X_1 - X_2$. Bestäm a så att $P(Y \geq a) = 0.905$.

A: ca 4.93

B: ca 4.27

C: ca -0.27

D: ca -0.93

Uppgift 6

Sannolikheten för vinst i ett visst spel är 0.25. Resultaten (dvs vinst/förlust) i successiva spel är oberoende av varandra. Låt X vara det antal spel som krävs tills två vinster har erhållits. Beräkna sannolikheten $P(X \geq 4)$.

A: 0.895

B: 0.844

C: 0.738

D: 0.105

Uppgift 7

En diskret stokastisk variabel X_i uppfyller

$$P(X_i = 1) = \frac{2}{5} - 2a, \quad P(X_i = 2) = \frac{2}{5} + a, \quad P(X_i = 3) = \frac{1}{5} + a,$$

där $i = 1, 2, 3$. Låt $x_1 = 2$, $x_2 = 2$ och $x_3 = 3$ vara observationer på de oberoende stokastiska variablerna X_1 , X_2 , och X_3 . Bestäm minsta-kvadrat-skattningen av a utgående från dessa tre observationer.

A: $2/45$

B: $1/9$

C: $2/15$

D: $8/45$

Uppgift 8

Låt $X_1 \in \text{Bin}(n_1, p)$ och $X_2 \in \text{Bin}(n_2, p)$ vara två oberoende stokastiska variabler. Parametern p är okänd. Vi betraktar två punktskattningar av p : p_{obs}^* och $\hat{p}_{\text{obs}}$ som ges av stickprovsvariablerna

$$p^* = \frac{X_1 + X_2}{n_1 + n_2} \text{ respektive } \hat{p} = \frac{X_1}{2n_1} + \frac{X_2}{2n_2}$$

Vilket påstående är sant givet att $n_1 = 25$ och $n_2 = 50$?

- A: p_{obs}^* är den effektivaste skattningen av p .
- B: $\hat{p}_{\text{obs}}$ är den effektivaste skattningen av p .
- C: Bägge skattningarna är lika effektiva.
- D: Man kan inte avgöra vilken av skattningarna som är effektivast, eftersom minst en av dem inte är väntevärdesriktig.

Uppgift 9

Låt X_i :na vara oberoende stokastiska variabler sådana att $X_i \in N(\mu, \sigma)$. $E(X_i) = \mu$, $D(X_i) = \sigma$, där μ och σ är okända.

Vi har ett stickprov med 25 observationer där vi fått $\bar{x} = 10$. Dessutom har vi fått ett ensidigt uppåt begränsat konfidensintervall med konfidensgrad 95% för σ^2 vars övre gräns är 15.65.

Vad blir övre gränsen för det ensidigt uppåt begränsade konfidensintervallet för μ som har konfidensgrad 95%?

- A: 10.69
- B: 11.03
- C: 11.35
- D: 11.63

Uppgift 10

Antag att $X \in Po(\mu)$. Vi vill testa $H_0 : \mu = 5$ mot alternativet $H_1 : \mu = 2$. Vi förkastar H_0 till förmån för H_1 om vi får ett utfall $x \leq 3$. Vi får en observation: $x = 1$.

Bestäm testets styrka.

- A: 0.040
- B: 0.265
- C: 0.406
- D: 0.857

Uppgift 11

Låt $\bar{x} = 137.0$, $\bar{y} = 208.0$, $s_x^2 = 92.5$, $s_y^2 = 163.0$, $n_x = 5$ och $n_y = 5$ vara givet och antag att $X_i \in N(\mu_x, \sigma)$ och $Y_i \in N(\mu_y, \sigma)$ samt att alla dessa stokastiska variabler är oberoende. Ange övre gränsen för det 95%-iga ensidigt uppåt begränsade konfidensintervallet $I_{\mu_x - \mu_y}$.

A: -50.2

B: -57.0

C: -57.7

D: -59.2

Uppgift 12

Datamängden 40, 60 anses vara utfall av stokastiska variabler $X_i; i = 1, 2$, där X_i :na antas vara Poissonfördelade med väntevärde μ . D.v.s. varje $X_i \in Po(\mu)$.

Datamängden 90, 90, 130 anses vara utfall av stokastiska variabler $Y_j; j = 1, 2, 3$, där Y_j :na antas vara Poissonfördelade med väntevärde 2μ . D.v.s. varje $Y_j \in Po(2\mu)$.

Alla stokastiska variablerna antas vara oberoende.

Vi skattar μ med $\mu_{obs}^* = \frac{\bar{x} + \bar{y}}{3}$

Bilda ett dubbelsidigt konfidensintervall med den approximativa konfidensgraden 90% för μ utgående från skattningen ovan och ange övre gränsen.

A: 55.3

B: 58.9

C: 62.9

D: 66.3

Var god vänd!

Del II

Uppgift 13

$X_1, X_2, \dots, X_n$, $n \geq 2$, är oberoende s.v. som alla är $N(0, 1)$. Låt

$$Y_1 = \frac{1}{n} \sum_{i=1}^n |X_i|, \quad Y_2 = \left| \frac{1}{n} \sum_{i=1}^n X_i \right|,$$

(a) Beräkna $E(Y_1)$. (5 p)

(b) Beräkna $E(Y_2)$ och ange $g(n)$ sådan att $E(Y_1) = g(n) \cdot E(Y_2)$. (5 p)

Uppgift 14

Livslängden för en viss slags elkomponent antas vara exponentialfördelad med väntevärdet μ timmar. Av n komponenter som kopplades in samtidigt visade det sig att k stycken fortfarande fungerade efter T timmar.

Bestäm Maximum-Likelihood-skattningen av μ då $n = 100, k = 90$ och $T = 1000$. (10 p)

Uppgift 15

Man ville undersöka om det finns en skillnad i opinion i en viss fråga bland sverigedemokrater (SD), moderater (M) och socialdemokrater (S). Man ställde därför en fråga till 40 (SD), 50 (M) och 50 (S) som de besvarade med "JA" eller "NEJ".

Resultatet blev att 26 (SD) svarade "JA", 21 (M) svarade "JA" och 27 (S) svarade "JA". Övriga (sammanlagt alltså 66 stycken) svarade "Nej".

Avgör om man med felrisken 5% kan säga att det föreligger en skillnad i opinion mellan de tre partiernas sympatisörer. (10 p)

Uppgift 16

(Fortsättning på Uppgift 6) Sannolikheten för vinst i ett visst spel är p . Resultaten (dvs vinst/förlust) i successiva spel är oberoende av varandra.

(a) Låt X vara det antal spel som krävs tills två vinster har erhållits. Bestäm sannolikhetsfunktionen för den stokastiska variabeln X . (5 p)

(b) Låt Y vara det antal spel som krävs tills tre vinster har erhållits. Bestäm sannolikhetsfunktionen för den stokastiska variabeln Y . (5 p)

Ledning: $1 + 2 + 3 + \dots + (n - 1) + n = \frac{n(n+1)}{2}$.

Lycka till!



Avd. Matematisk statistik

KTH Matematik

LÖSNINGSFÖRSLAG TENTAMEN I SF1922 SANNOLIKHETSTEORI OCH STATISTIK,
MÅNDAG 26 MAJ 2025 KL 8.00–13.00.

Del I, Svar.

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

Del I, Lösningsförslag.**Uppgift 1**

Påstående A och B är korrekta, enl definitioner. Det felaktiga påståendet står i C (rita Venn-diagram för C och D , för att se att det som står i D är ett korrekt påstående, och det som står i C - ett felaktigt påstående).

Svaret är C.

Uppgift 2

Vi har:

$$\begin{aligned} 1 &= \iint_{\mathbb{R}^2} f_{X,Y}(x,y) dx dy = \left| \text{rita integrations area: } x \geq 0, y \geq 0, x + y \leq 1 \right| = \\ &= \int_0^1 dx \left(\int_0^{1-x} cx dy \right) = \int_0^1 cx(1-x) dx = c \left(\frac{x^2}{2} - \frac{x^3}{3} \right) \Big|_{x=0}^{x=1} = c \left(\frac{1}{2} - \frac{1}{3} \right) = \frac{c}{6}, \end{aligned}$$

vilket ger alternativ A.

Uppgift 3

Vi har:

$$8 = V(X + 3Y) = V(X) + 9V(Y) + 2C(X, 3Y) = 2 + 18 + 6C(X, Y).$$

Dvs.

$$C(X, Y) = -12/6 = -2.$$

Därmed,

$$\rho(X, Y) = \frac{C(X, Y)}{D(X)D(Y)} = \frac{-2}{\sqrt{2}\sqrt{2}} = -1$$

Svaret är C.

Uppgift 4

Vi söker M sådan att:

$$F_{\xi}(M) = 0.5,$$

där $\xi \in \text{Exp}(\lambda)$, $\lambda = 1/200$. Dvs. vi har

$$F_{\xi}(M) = 1 - e^{-\lambda M} = 0.5,$$

eller

$$e^{-\lambda M} = 0.5.$$

När vi löser denna ekvation m.a.p. på M får vi:

$$M = -(1/\lambda) \ln(0.5) = \left| \lambda = 1/200 \right| = (-200) \ln(0.5) \approx 138.63 \approx 139$$

Svaret är C.

Uppgift 5

Det är känt att $X_1 \in N(2, 1)$ och $X_2 \in N(0, 2)$ är oberoende.

Följaktligen, s.v. $Y = X_1 - X_2 \in N(2, \sqrt{5})$,

dvs $\mu_Y = \mu_{X_1} - \mu_{X_2} = 2 - 0 = 2$, och $\sigma_Y = \sqrt{\sigma_{X_1}^2 + \sigma_{X_2}^2} = \sqrt{1^2 + 2^2} = \sqrt{5}$.

Därför, s.v. $Z = \frac{Y-2}{\sqrt{5}} \in N(0, 1)$.

Vidare, enligt Tabell 1 har vi för $Z \in N(0, 1)$:

$$P(Z \leq 1.31) = 0.9049 \simeq 0.905.$$

Detta, enligt symmetrin, leder till

$$P(Z \geq -1.31) \simeq 0.905.$$

Eftersom $Z = \frac{Y-2}{\sqrt{5}}$, får vi nu:

$$P\left(\frac{Y-2}{\sqrt{5}} \geq -1.31\right) \simeq 0.905,$$

eller

$$P(Y \geq 2 - 1.31\sqrt{5}) \simeq 0.905.$$

Därmed den sökta $a = 2 - 1.31\sqrt{5} \simeq -0.93$.

Svaret är D.

Uppgift 6

Låt s.v. ξ_1 vara det antal spel som krävs tills en vinst har erhållits, dvs $\xi_1 \in ffg(0.25)$. Låt ξ_2 vara en annan s.v. som är oberoende av ξ_1 och som har samma fördelning som ξ_1 , dvs $\xi_2 \in ffg(0.25)$ också.

Vi har: $X = \xi_1 + \xi_2$. Vidare,

$$P(X \geq 4) = 1 - (P(X = 2) + P(X = 3)).$$

För $P(X = 2)$, erhöles

$$\begin{aligned} P(X = 2) &= P(\xi_1 + \xi_2 = 2) = P(\{\xi_1 = 1\} \cap \{\xi_2 = 1\}) = |\text{oberoendet}| = \\ &= P(\xi_1 = 1) \cdot P(\xi_2 = 1) = (0.25)(0.25) = 0.25^2 \end{aligned}$$

För $P(X = 3)$, erhöles

$$\begin{aligned} P(X = 3) &= P(\xi_1 + \xi_2 = 3) = P(\{\xi_1 = 1\} \cap \{\xi_2 = 2\}) + P(\{\xi_1 = 2\} \cap \{\xi_2 = 1\}) = |\text{oberoendet}| = \\ &= P(\xi_1 = 1) \cdot P(\xi_2 = 2) + P(\xi_1 = 2) \cdot P(\xi_2 = 1) = (0.25)(0.25 \cdot 0.75) + (0.25 \cdot 0.75)(0.25) = 2 \cdot 0.25^2 \cdot (0.75). \end{aligned}$$

Därmed, $P(X \geq 4) = 1 - (P(X = 2) + P(X = 3)) = 1 - (0.25^2 + 2 \cdot (0.75) \cdot 0.25^2) = 0.84375 \simeq 0.844$

Svaret är B.

Uppgift 7

$$P(X_i = 1) = \frac{2}{5} - 2a, \quad P(X_i = 2) = \frac{2}{5} + a, \quad P(X_i = 3) = \frac{1}{5} + a,$$

$$E(X) = \sum_{j=1}^3 j p_X(j) = \frac{2}{5} - 2a + \frac{4}{5} + 2a + \frac{3}{5} + 3a = \frac{9}{5} + 3a$$

MK- skattningen av a är det a som minimerar Q i § 9.2 i F.S.

$$Q = \sum_{j=1}^3 (x_j - E(X))^2 = (2 - (\frac{9}{5} + 3a))^2 + (2 - (\frac{9}{5} + 3a))^2 + (3 - (\frac{9}{5} + 3a))^2$$

$$\frac{dQ}{da} = -2(2 - (\frac{9}{5} + 3a))3 - 2(2 - (\frac{9}{5} + 3a))3 - 2(3 - (\frac{9}{5} + 3a))3 = 0$$

$$\Rightarrow 4 - \frac{18}{5} - 6a + 3 - \frac{9}{5} - 3a = 0 \Rightarrow 7 - \frac{27}{5} = 9a \Rightarrow a = \frac{8}{45}$$

$$\text{Svar: } a = \frac{8}{45}$$

Uppgift 8

$$p^* = \frac{X_1 + X_2}{n_1 + n_2} \text{ respektive } \hat{p} = \frac{X_1}{2n_1} + \frac{X_2}{2n_2}$$

$$E(p^*) = E\left(\frac{X_1 + X_2}{n_1 + n_2}\right) = \frac{1}{n_1 + n_2}(E(X_1) + E(X_2)) = \frac{1}{n_1 + n_2}(n_1 p + n_2 p) = p$$

$$E(\hat{p}) = E\left(\frac{X_1}{2n_1}\right) + E\left(\frac{X_2}{2n_2}\right) = \frac{n_1 p}{2n_1} + \frac{n_2 p}{2n_2} = p$$

Således är både p^* och $\hat{p}$ väntevärdesriktiga skattningar av p .

$$V(p^*) = V\left(\frac{X_1 + X_2}{n_1 + n_2}\right) = \frac{1}{(n_1 + n_2)^2}(V(X_1) + V(X_2)) = \frac{1}{(n_1 + n_2)^2}(n_1 p(1-p) + n_2 p(1-p)) = \frac{p(1-p)}{n_1 + n_2}$$

$$V(\hat{p}) = V\left(\frac{X_1}{2n_1}\right) + V\left(\frac{X_2}{2n_2}\right) = \frac{n_1 p(1-p)}{(4n_1^2)} + \frac{n_2 p(1-p)}{(4n_2^2)} = \frac{p(1-p)}{4}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

Med $n_1 = 25$ och $n_2 = 50$ fås

$$V(p^*) = \frac{p(1-p)}{75} = \frac{3p(1-p)}{225} \text{ och } V(\hat{p}) = \frac{p(1-p)}{4}\left(\frac{1}{25} + \frac{1}{50}\right) = \frac{3p(1-p)}{200}$$

$V(p^*) < V(\hat{p})$ Alltså gäller att p_{obs}^* är en effektivare skattning av p än $\hat{p}_{obs}$.

Uppgift 9

Det ensidigt uppåt begränsade konfidensintervallet för σ^2 fås m.h.a. § 12.4 och § 11.1b till

$$I_{\sigma^2} = \left(0, \frac{(n-1)s^2}{\chi_{1-\alpha}^2(n-1)}\right) = \left(0, \frac{24s^2}{\chi_{0.95}^2(24)}\right) = \left(0, \frac{24s^2}{13.8}\right) = \text{enl uppg.} = (0, 15.65)$$

Detta ger att $s = 3$.

Med hjälp av §12.2 och §11.1d fås övre gränsen till det ensidigt uppåt begränsade konfidensintervallet för μ till

$$I_{\mu} = \bar{x} + \frac{s}{\sqrt{n}}t_{\alpha}(n-1) = 10 + \frac{3}{\sqrt{25}}t_{0.05}(24) = 10 + 0.6 \cdot 1.71 = 11.03$$

Uppgift 10

Styrkan hos testet är sannolikheten att förkasta nollhypotesen om mothypotesen är sann. Dvs.

$$P(\text{Förkasta } H_0) \text{ om } H_1 \text{ är sann} = P(X \leq 3) \text{ om } \mu = 2.$$

Vi tittar i tab 5 som ger $P(X \leq 3)$ om $X \in Po(2)$ och får svaret 0.857

Uppgift 11

Låt $\bar{x} = 137.0$, $\bar{y} = 208.0$, $s_x^2 = 92.5$, $s_y^2 = 163.0$, $n_x = 5$ och $n_y = 5$

Ett ensidigt uppåt begränsat konfidensintervall för skillnaden mellan två stickprovs väntevärden där vi antar att stickproven har samma okända varians fås m.h.a. §12.2 och § 11.2 till

$$I_{\mu_x - \mu_y} = \left(-\infty, \bar{x} - \bar{y} + s \cdot \sqrt{\frac{1}{n_x} + \frac{1}{n_y}} \cdot t_{\alpha}(n_x + n_y - 2)\right)$$

där

$$s^2 = \frac{(n_x - 1) \cdot s_x^2 + (n_y - 1) \cdot s_y^2}{n_x + n_y - 2} = \frac{(5 - 1) \cdot 92.5 + (5 - 1) \cdot 163.0}{5 + 5 - 2} = 127.75$$

Den övre gränsen i konfidensintervallet kan nu skrivas

$$I_{\mu_x - \mu_y} = 137.0 - 208.0 + \sqrt{127.75} \cdot \sqrt{\frac{1}{5} + \frac{1}{5}} \cdot t_{0.05}(8) = -71 + \sqrt{127.75} \cdot \sqrt{\frac{2}{5}} \cdot 1.86 = -71 + 13.3 = -57.7$$

Vilket ger att övre gränsen blir -57.7 .

Uppgift 12

Vi har Poisson-fördelade data och alla utfall är här större än 15 så därför kan vi enligt §5 i F.S. göra Normalapproximation. Då kan vi bilda ett konfidesintervall med approximativ konfidensgrad m.h.a. §12.3 i F.S.

$$I_\mu = \mu_{obs}^* \pm D_{obs}^*(\mu^*)\lambda_{\frac{\alpha}{2}}$$

$$\mu_{obs}^* = \frac{\bar{x} + \bar{y}}{3}$$

$$V(\mu^*) = V\left(\frac{\bar{X} + \bar{Y}}{3}\right) = \frac{1}{9}V(\bar{X} + \bar{Y}) = \text{ober} = \frac{1}{9}[V(\bar{X}) + V(\bar{Y})] = \text{ober} = \frac{1}{9}\left[\frac{V(X_i)}{2} + \frac{V(Y_j)}{3}\right] =$$

$$= \frac{1}{9}\left[\frac{\mu}{2} + \frac{2\mu}{3}\right] = \frac{1}{9} \cdot \frac{7\mu}{6} = \frac{7\mu}{54} \Rightarrow D_{obs}^*(\mu_{obs}^*) = \sqrt{\frac{7\mu_{obs}^*}{54}} = \sqrt{\frac{7}{54}\left(\frac{\bar{x} + \bar{y}}{3}\right)}$$

$$\bar{x} = \frac{40 + 60}{2} = 50 \text{ och } \bar{y} = \frac{90 + 90 + 130}{3} = \frac{310}{3} \Rightarrow \frac{\bar{x} + \bar{y}}{3} = \frac{460}{9}$$

$$\text{Nu fås } I_\mu = \mu_{obs}^* \pm D_{obs}^*(\mu^*)\lambda_{\frac{\alpha}{2}} = \frac{460}{9} \pm \sqrt{\frac{7}{54}\left(\frac{460}{9}\right)}\lambda_{0.05} = \frac{460}{9} \pm \sqrt{\frac{7}{54}\left(\frac{460}{9}\right)}1.6449$$

Övre gränsen blir då 55.3

Del II, Lösningsförslag.**Uppgift 13**

(a) Vi har:

$$E(Y_1) = |\text{linjaritet}| = \frac{1}{n} \sum_{i=1}^n E(|X_i|) = |\text{samma fördelning av } X_i| = E(|X_1|).$$

Vidare, eftersom $X_1 \in N(0, 1)$,

$$\begin{aligned} E(Y_1) = E(|X_1|) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |y| e^{-\frac{y^2}{2}} dy = |\text{symmetri}| = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} y e^{-\frac{y^2}{2}} dy = \\ &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-\frac{y^2}{2}} d\left(\frac{y^2}{2}\right) = \sqrt{\frac{2}{\pi}} \cdot \left[-e^{-\frac{y^2}{2}} \right]_0^{\infty} = \sqrt{\frac{2}{\pi}} \end{aligned}$$

(b) Vi söker $E(Y_2) = E\left|\frac{1}{n} \sum_{i=1}^n X_i\right| = E|\bar{X}|$, där $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.
Eftersom alla $X_i \in N(0, 1)$, $i = 1, 2, \dots, n$ och oberoende, har vi:

$$\bar{X} \in N\left(0, \frac{1}{\sqrt{n}}\right)$$

Därmed,

$$\begin{aligned} E(Y_2) = E|\bar{X}| &= \frac{\sqrt{n}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |y| e^{-\frac{n}{2}y^2} dy = |\text{symmetri}| = \frac{2\sqrt{n}}{\sqrt{2\pi}} \int_0^{\infty} y e^{-\frac{n}{2}y^2} dy = \\ &= \frac{2\sqrt{n}}{n\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{n}{2}y^2} d\left(\frac{n}{2}y^2\right) = \\ &= \frac{2\sqrt{n}}{n\sqrt{2\pi}} \cdot \left[-e^{-\frac{n}{2}y^2} \right]_0^{\infty} = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sqrt{n}} = \\ &= \frac{E(Y_1)}{g(n)}, \end{aligned}$$

där $g(n) = \sqrt{n}$.

Uppgift 14

Låt den stokastiska variabeln X vara det antal komponenter som har en livslängd $> T$.

Då gäller att $X \in \text{Bin}(n, p)$.

Låt den stokastiska Y vara livslängden hos en enskild komponent.

Då gäller att $Y \in \text{Exp}(\lambda) = \text{Exp}(\frac{1}{\mu})$.

$$p = P(Y > T) = \int_T^\infty \frac{1}{\mu} e^{-\frac{t}{\mu}} dt = e^{-\frac{T}{\mu}}$$

se §9.1 i F.S. $L(\mu) = p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} (e^{-\frac{T}{\mu}})^k (1 - e^{-\frac{T}{\mu}})^{n-k}$

Maximum för $L(\mu)$ finns där maximum för $\ln L(\mu)$ finns.

$$\ln L(\mu) = \ln \binom{n}{k} + \ln(e^{-\frac{T}{\mu}})^k + \ln(1 - e^{-\frac{T}{\mu}})^{n-k} = \ln \binom{n}{k} - \frac{kT}{\mu} + (n-k) \ln(1 - e^{-\frac{T}{\mu}})$$

$$\frac{d}{d\mu} \ln L(\mu) = \frac{kT}{\mu^2} + (n-k) \frac{-\frac{T}{\mu^2} e^{-\frac{T}{\mu}}}{1 - e^{-\frac{T}{\mu}}} = 0$$

$$\Rightarrow k = \frac{(n-k)e^{-\frac{T}{\mu}}}{1 - e^{-\frac{T}{\mu}}} = \frac{(n-k)a}{1-a}$$

$$\frac{k}{n-k} = \frac{a}{1-a} \Rightarrow k - ak = an - ak \Rightarrow k = an \Rightarrow a = \frac{k}{n} = e^{-\frac{T}{\mu}}$$

$$\ln\left(\frac{k}{n}\right) = -\frac{T}{\mu} \Rightarrow \mu_{obsMK}^* = \frac{T}{\ln\left(\frac{n}{k}\right)} = \frac{1000}{\ln\left(\frac{100}{90}\right)} = 9491.2 \text{ timmar}$$

Uppgift 15

	<i>JA</i>	<i>NEJ</i>	Totalt
SD	26	14	40
M	21	29	50
S	27	23	50
Totalt	74	66	140

Homogenitetstest med $r = 2$ (antal kategorier) och $s = 3$ (antal serier). Vi får enligt FS 14.3 teststorheten

$$Q = \sum_{i=1}^3 \sum_{j=1}^2 \frac{(x_{ij} - \frac{n_i m_j}{N})^2}{\frac{n_i m_j}{N}}$$

där $n_1 = 40$, $n_2 = 50$, $n_3 = 50$, $m_1 = 26 + 21 + 27 = 74$, $m_2 = 66$, $N = 140$, $x_{11} = 26$, $x_{12} = 40 - 26 = 14$, $x_{21} = 21$, $x_{22} = 29$, $x_{31} = 27$, $x_{32} = 23$.

$$\frac{n_1 m_2}{N} = \frac{40 \cdot 66}{140} \geq 5 \text{ så då är alla } \frac{n_i m_j}{N} \geq 5 \text{ och då är det okej med } \chi^2\text{-test.}$$

Detta ger $Q = 4.76$ som är utfall av approximativ $\chi^2((r-1)(s-1)) = \chi^2(2)$ -fördelning. Eftersom $\chi_{0.05}^2(2) = 5.99$ kan H_0 : 'samma andel i alla tre partierna' ej förkastas ty $Q = 4.76 < 5.99 = \chi_{0.05}^2(2)$. Ingen skillnad alltså visad!

Uppgift 16

(Fortsättning på Uppgift 6) Låt p vara sannolikheten för vinst i ett visst spel.

(a) Låt s.v. ξ_1 vara det antal spel som krävs tills en vinst har erhållits, dvs $\xi_1 \in ffg(p)$:

$$p_{\xi_1}(j) = P(\xi_1 = j) = p(1-p)^{j-1}, \quad j = 1, 2, 3, \dots$$

Låt ξ_2 vara en annan s.v. som är oberoende av ξ_1 och som har samma fördelning som ξ_1 , dvs $\xi_2 \in ffg(p)$ också:

$$p_{\xi_2}(j) = P(\xi_2 = j) = p(1-p)^{j-1}, \quad j = 1, 2, 3, \dots$$

Vi har: $X = \xi_1 + \xi_2$.

Faltningsformeln för oberoende diskreta s.v. ξ_1 och ξ_2 ger oss sannolikhetsfunktionen för s.v. $X = \xi_1 + \xi_2$, dvs.

$$p_X(k) = P(X = k) = \sum_{j=1}^{k-1} p_{\xi_1}(j)p_{\xi_2}(k-j),$$

där $k = 2, 3, 4, \dots$ och begränsningen $k-1$ i summan kommer från det villkoret att $k-j \geq 1$ dvs $j \leq k-1$. Vidare, har vi eftersom ξ_1 och ξ_2 är oberoende och både $ffg(p)$ -fördelade:

$$\begin{aligned} p_X(k) &= \sum_{j=1}^{k-1} p_{\xi_1}(j)p_{\xi_2}(k-j) = \sum_{j=1}^{k-1} \left(p(1-p)^{j-1} \right) \left(p(1-p)^{k-j-1} \right) = \\ &= \sum_{j=1}^{k-1} p^2(1-p)^{k-2} = |ej \text{ beror av } j| = (k-1)p^2(1-p)^{k-2}. \end{aligned}$$

Alltså, sannolikhetsfunktionen för s.v. X är

$$p_X(k) = P(X = k) = (k-1)p^2(1-p)^{k-2}, \quad k = 2, 3, 4, \dots$$

OBS! Man kan nu dubbelkolla mha $p_X(k)$ att svaret i Uppgift 6 stämmer.

(b) Låt en annan s.v. ξ_3 vara också det antal spel som krävs tills en vinst har erhållits, dvs $\xi_3 \in \text{ffg}(p)$:

$$p_{\xi_3}(j) = P(\xi_3 = j) = p(1-p)^{j-1}, \quad j = 1, 2, 3, \dots$$

Låt ξ_3 vara oberoende av ξ_1 och ξ_2 från (a).

Vi har isf att s.v. Y är:

$$Y = \xi_1 + \xi_2 + \xi_3 = X + \xi_3.$$

Eftersom alla tre ξ_i är oberoende stokastiska variabler, så är oberoende även s.v. $X = \xi_1 + \xi_2$ och ξ_3 .

Vidare, på samma sätt som i (a), faltningsformeln för oberoende diskreta s.v. X och ξ_3 ger oss sannolikhetsfunktionen för s.v. Y , dvs.

$$p_Y(k) = P(Y = k) = \sum_{j=2}^{k-1} p_X(j) p_{\xi_3}(k-j),$$

där $k = 3, 4, 5, \dots$ och begränsningar $k-1$ i summan kommer från det villkoret att $k-j \geq 1$ dvs $j \leq k-1$, samt $j \geq 2$ kommer från att $p_X(j)$ är definierad fr.o.m. $j = 2$.

Därmed, har vi enl faltningen:

$$\begin{aligned} p_Y(k) &= \sum_{j=2}^{k-1} p_X(j) p_{\xi_3}(k-j) = \sum_{j=2}^{k-1} \left((j-1)p^2(1-p)^{j-2} \right) \left(p(1-p)^{k-j-1} \right) = \\ &= p^3(1-p)^{k-3} \sum_{j=2}^{k-1} (j-1) = |j-1 = m| = p^3(1-p)^{k-3} \sum_{m=1}^{k-2} m = \\ &= | \text{enl Ledning} | = \\ &= \frac{(k-2)(k-1)}{2} p^3 (1-p)^{k-3}. \end{aligned}$$

Alltså, sannolikhetsfunktionen för s.v. Y är

$$p_Y(k) = P(Y = k) = \frac{(k-2)(k-1)}{2} p^3 (1-p)^{k-3}, \quad k = 3, 4, 5, \dots$$