



KTH Matematik

Avd. Matematisk statistik

TENTAMEN I SF1910,SF1925/SF1917,SF1919
TILLÄMPAD STATISTIK/SANNOLIKHETSTEORI OCH STATISTIK
TISDAG 7 APRIL 2026 KL 08.00–13.00.

Kursledare: 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 A och B vara två oberoende händelser. Följande sannolikheter är kända

$$P(A \cup B) = 0.92, \quad P(A \cup B^*) = 0.88, \quad P(A^* \cup B) = 0.68.$$

Bestäm $P(A \cap B)$.

A: 0.36

B: 0.48

C: 0.52

D: 0.64

Uppgift 2

Den stokastiska variabeln X har fördelningsfunktion

$$F_X(x) = \begin{cases} 0, & x \leq -5 \\ \frac{10+2x}{20}, & \text{om } -5 \leq x \leq 5 \\ 1, & x \geq 5 \end{cases}$$

$$Y = X^2.$$

Beräkna $P(Y \leq 4)$.

- A: 0.16
- B: 0.20
- C: 0.32
- D: 0.40

Uppgift 3

Antag att X_1, \dots, X_4 utgör ett stickprov. $X_i \in N(3, \sigma)$. $\sigma = 4$.

Bestäm a så att $P(|\bar{X} - 3| > a) = 0.05$.

- A: 1.65
- B: 1.96
- C: 3.29
- D: 3.92

Uppgift 4

Den stokastiska variabeln X har täthetsfunktionen

$$f_X(x) = \begin{cases} 12x^2(1-x), & \text{om } 0 \leq x \leq 1 \\ 0, & \text{annars} \end{cases}$$

Beräkna variansen $V(X)$.

- A: 0.02
- B: 0.04
- C: 0.08
- D: 0.16

Uppgift 5

De oberoende stokastiska variablerna X och Y är sådana att $p_X(1) = \frac{1}{3}$, och $p_X(2) = \frac{2}{3}$, samt $p_Y(1) = \frac{1}{6}$ och $p_Y(2) = \frac{5}{6}$. Den stokastiska variabeln $Z = X + Y$.

Beräkna variansen $V(Z)$.

A: 0.16

B: 0.25

C: 0.36

D: 0.49

Uppgift 6

De stokastiska variablerna X och Y har standardavvikelserna $D(X) = 8$ och $D(Y) = 6$. De är beroende på så sätt att kovariansen är $C(X, Y) = 30$.

Beräkna standardavvikelsen $D(2X - 3Y + 7)$.

A: 14.8

B: 22.8

C: 23.6

D: 30.7

Uppgift 7

Antag att X är fördelat enligt $\text{Exp}(\theta)$ medan Y är fördelat enligt $\text{Exp}(2\theta)$.

Beräkna Minsta-Kvadrat-skattningen för θ givet $x = 2$ och $y = 4$.

A: 0.15

B: 0.20

C: 5.0

D: 6.8

Uppgift 8

Låt $\bar{x} = 237.0$, $\bar{y} = 108.0$, $s_x = 9.2$, $s_y = 12.6$, $n_x = 5$ 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 99%-iga dubbelsidiga konfidensintervallet $I_{\mu_x - \mu_y}$ för $\mu_x - \mu_y$.

A: 146.19

B: 148.35

C: 152.44

D: 159.70

Uppgift 9

Låt X vara fördelad enligt $\text{Bin}(177, p)$ och antag att vi observerar $x = 95$. Ange nedre gränsen för det ensidiga konfidensintervallet för p om den approximativa konfidensgraden är 95% och vi skattar p med $p_{obs}^* = x/177$.

A: 0.446

B: 0.463

C: 0.475

D: 0.489

Uppgift 10

Vi antar att $X \in \text{Po}(\mu)$ och testar nollhypotesen $H_0 : \mu = 10$ mot mothypotesen $H_1 : \mu = 5$ och får utfallet $x = 4$. Vi förkastar H_0 till förmån för H_1 om vi får att $x \leq 6$.

Beräkna styrkan hos testet.

A: 0.03

B: 0.13

C: 0.44

D: 0.76

Uppgift 11

I en marknadsundersökning ville man jämföra olika kundgruppers preferenser för två konkurrerande produkter A och B. 100 testpersoner i åldersintervallet 18–30 år fick prova de två produkterna och avgöra om de föredrog produkt A eller produkt B eller ifall de inte ansåg att det förelåg någon skillnad mellan produkterna. Resultaten sammanfattas av tabellen

	Föredrar A	Föredrar B	Ingen skillnad	
Åldersgrupp 18–30 år	23	47	30	Totalt: 100

Samma undersökning gjordes på 50 personer i åldersintervallet 30–65 år. Resultaten sammanfattas av tabellen

	Föredrar A	Föredrar B	Ingen skillnad	
Åldersgrupp 30–65 år	16	13	21	Totalt: 50

Vi beräknar teststorheten Q och får $Q = 6.1$

- A: Skillnaden mellan de två åldersgruppernas produktval i detta fall är varken signifikant på risknivån 1% eller risknivån 5%.
- B: Skillnaden mellan de två åldersgruppernas produktval i detta fall är signifikant både på risknivån 1% och risknivån 5%.
- C: Skillnaden mellan de två åldersgruppernas produktval i detta fall är signifikant på risknivån 1%, men inte på risknivån 5%.
- D: Skillnaden mellan de två åldersgruppernas produktval i detta fall är signifikant på risknivån 5%, men inte på risknivån 1%.

Uppgift 12

I en studie presenteras sambandet mellan tryckhållfasthet x och permeabilitet y för olika blandningar av betong.

En sammanfattning av dessa data, som omfattar $n = 14$ par (x_i, y_i) av observationer, är

$$\sum_{i=1}^{14} y_i = 572 \qquad \sum_{i=1}^{14} y_i^2 = 23530$$

$$\sum_{i=1}^{14} x_i = 43 \qquad \sum_{i=1}^{14} x_i^2 = 157.42$$

$$\sum_{i=1}^{14} x_i y_i = 1697.80$$

Vi antar att $(x_1, y_1), \dots, (x_{14}, y_{14})$ är relaterade till varandra enligt en linjär regressionsmodell dvs. som utfall av $(x_1, Y_1), \dots, (x_{14}, Y_{14})$, där varje Y_i är normalfördelad $N(\alpha + \beta x_i, \sigma)$.

Skatta regressionslinjen och använd sedan den skattade regressionslinjen för att skatta den förväntade permeabiliteten y när tryckhållfastheten $x = 4.3$.

A: skattningen av y är ca 38

B: skattningen av y är ca 48

C: skattningen av y är ca 58

D: skattningen av y är ca 68

Del II

Uppgift 13

Vid en tillverkningsprocess kontrolleras de tillverkade enheterna i en datorstyrd sensor. Härvid klassificeras defekta enheter som defekta med sannolikheten 0.9 och som korrekta med sannolikheten 0.1. Vidare klassificeras korrekta enheter som korrekta med sannolikheten 0.85 och som defekta med sannolikheten 0.15. Vad är den betingade sannolikheten att en enhet är defekt givet att den klassificerats som defekt, om processens felsannolikhet är 0.1? (10 p)

Uppgift 14

a) En tjuv som tar ditt bankomatkort försöker att knäcka koden. Koden består av fyra siffror. Efter tre misslyckade försök så tar bankomaten hand om kortet. Tjuven väljer kod helt slumpmässigt, dock inte samma kod två gånger på ett och samma kort. Beräkna sannolikheten att tjuven knäcker koden innan kortet försvinner in i bankomaten efter tre försök. (4 p)

b) Antag att det av olika tjuvar stjäls 10000 bankomatkort under ett år. För varje kort så görs det försök, att på samma sätt som ovan, att knäcka koden. Tjuvarna utbyter inte information sinsemellan om vilka koder som de har använt för att försöka knäcka koden på "deras" kort. Vad är sannolikheten att koden knäcks på tre eller flera kort? Välmotiverad approximation får användas. Om du inte har svarat på a) så får du anta att svaret där skall vara 0.0001. (6 p)

Uppgift 15

Två företag, A och B, har telefon-"support", och man ville uppskatta skillnaden i väntetid i telefon innan man kommer fram till "supporten" mellan de två företagen. Av 420 telefonsamtal till företag A blev genomsnittliga väntetiden 26.0 minuter, av 376 telefonsamtal till B blev genomsnittliga väntetiden 31.6 minuter. Antag att bägge företagen har exponentialfördelade väntetider, och bestäm ett approximativt konfidensintervall med konfidensgrad approximativt 95% för skillnaden i väntevärden (av väntetiderna) $\mu_B - \mu_A$. Ange också om någon slutsats kan dras med approximativ felrisk 5% angående vilket företag som har den längsta förväntade väntetiden. Svaret skall alltså både innehålla ett konfidensintervall och en sådan slutsats. (10 p)

Uppgift 16

Antag allmänt att man har två oberoende stickprov på $N(\mu, \sigma)$ med lika många, n , observationer i båda. Vi antar här att σ är känt. Bilda från stickproven var sitt 95% konfidensintervall för μ . Vad är sannolikheten att konfidensintervallen inte skär varandra, d.v.s inte har någon punkt gemensam? (10 p)

Lycka till!



Avd. Matematisk statistik

KTH Matematik

LÖSNINGSFÖRSLAG TENTAMEN I SF1910/SF1917/SF1919/SF1925
SANNOLIKHETSTEORI OCH STATISTIK,
TISDAG 7 APRIL 2026 KL 8.00–13.00.

Del I, Svar.

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

Del I, Lösningsförslag.**Uppgift 1**

$$P(A \cup B) = 0.92, \quad P(A \cup B^*) = 0.88, \quad P(A^* \cup B) = 0.68.$$

Bestäm $P(A \cap B)$.

Vi börjar med att ta fram de tre komplementära händelserna till de tre kända. Då får vi $P(A^* \cap B^*) = 0.08$, $P(A^* \cap B) = 0.12$, $P(A \cap B^*) = 0.32$.

$$P(A^*) = P(A^* \cap B^*) + P(A^* \cap B) = 0.08 + 0.12 = 0.20.$$

$$P(B^*) = P(A^* \cap B^*) + P(A \cap B^*) = 0.08 + 0.32 = 0.40.$$

Då får vi att $P(A \cap B) = [ober] = P(A)P(B) = 0.8 \cdot 0.6 = 0.48$

Uppgift 2

Den stokastiska variabeln X har fördelningsfunktionenfunktionen

$$F_X(x) = \begin{cases} 0, & x \leq -5 \\ \frac{10+2x}{20}, & \text{om } -5 \leq x \leq 5 \\ 1, & x \geq 5 \end{cases}$$

$$Y = X^2.$$

$$P(Y \leq 4) = P(X^2 \leq 4) = P(-2 \leq X \leq 2) = P(X \leq 2) - P(X \leq -2) = F_X(2) - F_X(-2) =$$

$$= \frac{10 + 2 \cdot 2}{20} - \frac{10 + 2 \cdot (-2)}{20} = \frac{14}{20} - \frac{6}{20} = 0.4$$

Uppgift 3

Antag att X_1, \dots, X_4 utgör ett stickprov. $X_i \in N(3, \sigma)$. $\sigma = 4$.

Bestäm a så att $P(|\bar{X} - 3| > a) = 0.05$.

Dvs. $P(|\bar{X} - 3| < a) = 0.95$

$$P(|\bar{X} - 3| < a) = P(-a < \bar{X} - 3 < a) = P(3 - a < \bar{X} < 3 + a) = 0.95$$

Dags att göra om till $N(0, 1)$

$E(\bar{X}) = E(X) = 3$. $D(\bar{X}) = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{4}} = 2$. Då får vi

$$P\left(\frac{3 - a - 3}{2} < \frac{\bar{X} - 3}{2} < \frac{3 + a - 3}{2}\right)$$

Nu har vi

$$Z = \frac{\bar{X} - 3}{2} \in N(0, 1)$$

Dvs.

$$P\left(\frac{-a}{2} < Z < \frac{a}{2}\right) = 0.95$$

Symmetri ger då att $P(Z > \frac{a}{2}) = 0.025$

Tabell 2 ger då att $\frac{a}{2} = \lambda_{0.025} = 1.96 \Rightarrow a = 3.92$

Uppgift 4

För att beräkna variansen för X använder vi beräkningsformeln

$$V(X) = E(X^2) - (E(X))^2$$

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 12x^3 (1 - x) dx = [12\frac{x^4}{4} - 12\frac{x^5}{5}]_0^1 = 12(\frac{1}{4} - \frac{1}{5}) = \frac{12}{20} = 0.6$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^1 12x^4 (1 - x) dx = [12\frac{x^5}{5} - 12\frac{x^6}{6}]_0^1 = 12(\frac{1}{5} - \frac{1}{6}) = \frac{12}{30} = 0.4$$

$$V(X) = E(X^2) - E^2(X) = 0.4 - 0.6^2 = 0.04$$

Uppgift 5

Vi inleder med att bestämma fördelningen för $Z = X + Y$. Vi använder faltningsformeln för oberoende stokastiska variabler

$$p_Z(k) = \sum_{i+j=k} p_X(i) p_Y(j).$$

Om man tänker efter inser man att $k = 2, 3, 4$.

$$p_Z(2) = p_X(1) p_Y(1) = \frac{1}{3} \cdot \frac{1}{6} = \frac{1}{18}.$$

$$p_Z(3) = p_X(1) p_Y(2) + p_X(2) p_Y(1) = \frac{1}{3} \cdot \frac{5}{6} + \frac{2}{3} \cdot \frac{1}{6} = \frac{7}{18}.$$

$$p_Z(4) = p_X(2) p_Y(2) = \frac{2}{3} \cdot \frac{5}{6} = \frac{10}{18}.$$

För att beräkna variansen för Z använder vi beräkningsformeln

$$V(Z) = E(Z^2) - (E(Z))^2$$

$$E(Z) = \sum_{\text{alla } k} k p_Z(k) = 2 \cdot \frac{1}{18} + 3 \cdot \frac{7}{18} + 4 \cdot \frac{10}{18} = \frac{63}{18} = \frac{7}{2}$$

$$E(Z^2) = \sum_{\text{alla } k} k^2 p_Z(k) = 2^2 \cdot \frac{1}{18} + 3^2 \cdot \frac{7}{18} + 4^2 \cdot \frac{10}{18} = \frac{227}{18}$$

$$V(Z) = E(Z^2) - (E(Z))^2 = \frac{227}{18} - \left(\frac{7}{2}\right)^2 = \frac{227}{18} - \frac{49}{4} = \frac{13}{36}$$

Då $\frac{13}{36} = 0.361$ är alltså $V(Z) = 0.36$.

Uppgift 6

$$\begin{aligned} V(2X - 3Y + 7) &= V(2X - 3Y) \\ &= 2^2 \cdot V(X) + (-3)^2 \cdot V(Y) + 2C(2X, -3Y) \\ &= 4 \cdot V(X) + 9 \cdot V(Y) + 2 \cdot 2 \cdot (-3)C(X, Y) \\ &= 4 \cdot 8^2 + 9 \cdot 6^2 + 2 \cdot 2 \cdot (-3) \cdot 30 \\ &= 220 \end{aligned}$$

Därmed blir

$$D(2X - 3Y + 7) = \sqrt{V(2X - 3Y + 7)} = \sqrt{220} = 14.8$$

Uppgift 7

Antag att X är fördelat enligt $Exp(\theta)$ medan Y är fördelat enligt $Exp(2\theta)$.

Beräkna Minsta-Kvadrat-skattningen för θ givet $x = 2$ och $y = 4$.

De båda stokastiska variablerna har olika varians, så här använder vi oss av formeln

$$Q = Q(\theta_1, \theta_2, \dots, \theta_k) = \sum_{i=1}^n w_i (x_i - \mu_i(\theta_1, \theta_2, \dots, \theta_k))^2 \text{ där } w_i = 1/\sigma_i^2.$$

Enligt §4 i F.S fås om $X \in Exp(a\theta)$ att $E(X) = D(X) = \frac{1}{a\theta}$

och om $Y \in Exp(b\theta)$ att $E(Y) = D(Y) = \frac{1}{b\theta}$

Detta leder till att

$$Q = (a\theta)^2 \left(x - \frac{1}{a\theta}\right)^2 + (b\theta)^2 \left(y - \frac{1}{b\theta}\right)^2 = (a\theta)^2 \left(x^2 + \frac{1}{a^2\theta^2} - \frac{2x}{a\theta}\right) + (b\theta)^2 \left(y^2 + \frac{1}{b^2\theta^2} - \frac{2y}{b\theta}\right)$$

Dvs

$$Q = a^2\theta^2 x^2 + 1 - 2xa\theta + b^2\theta^2 y^2 + 1 - 2yb\theta$$

$$\frac{dQ}{d\theta} = 2a^2\theta x^2 - 2xa + 2b^2\theta y^2 - 2yb = 0$$

Detta ger

$$\theta_{obs_{MK}}^* = \frac{ax + by}{a^2x^2 + b^2y^2} = [\text{i vårt fall}] = \frac{1 \cdot 2 + 2 \cdot 4}{1^2 \cdot 2^2 + 2^2 \cdot 4^2} = \frac{10}{68} = 0.15$$

Uppgift 8

Låt $\bar{x} = 237.0$, $\bar{y} = 108.0$, $s_x = 9.2$, $s_y = 12.6$, $n_x = 5$ 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 99%-iga dubbelsidiga konfidensintervallet $I_{\mu_x - \mu_y}$ för $\mu_x - \mu_y$.

Ett tvåsidigt 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

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

där

$$s^2 = \frac{(n_x - 1) \cdot s_x^2 + (n_y - 1) \cdot s_y^2}{n_x + n_y - 2} = \frac{(5 - 1) \cdot 9.2^2 + (5 - 1) \cdot 12.6^2}{5 + 5 - 2} = 121.7$$

Intervallet $I_{\mu_x - \mu_y}$ nu skrivs

$$I_{\mu_x - \mu_y} = 237 - 108 \pm \sqrt{121.7} \cdot \sqrt{\frac{1}{5} + \frac{1}{5}} \cdot t_{0.005}(8) = 129 \pm \sqrt{\frac{121.7 \cdot 2}{5}} \cdot 3.36 = 129 \pm 23.44$$

Viket ger att övre gränsen blir 152.44.

Uppgift 9

$X \in \text{Bin}(n, p) = \text{Bin}(177, p)$ Vi får då konfidensintervallets nedre gräns till

$$\begin{aligned} I_p &= p_{obs}^* - \sqrt{\frac{p_{obs}^* \cdot (1 - p_{obs}^*)}{n}} \cdot \lambda_\alpha = \frac{95}{177} - \sqrt{\frac{\frac{95}{177} \cdot (1 - \frac{95}{177})}{177}} \cdot \lambda_{0.05} = \\ &= \frac{95}{177} - \sqrt{\frac{\frac{95}{177} \cdot (1 - \frac{95}{177})}{177}} \cdot 1.6449 = 0.475 \end{aligned}$$

Uppgift 10

Styrkan hos testet är sannolikheten att förkasta H_0 om H_1 är sann $= P(X \leq 6)$ om $X \in \text{Po}(5)$. Tab 5 ger då svaret 0.76.

Uppgift 11

Homogenitetstest. Nollhypotesen är att det inte är någon skillnad i preferenser mellan de två åldersgrupperna. Vi förkastar nollhypotesen för stora värden på Q . som om nollhypotesen är sann är ett utfall från en (approximativt) $\chi^2((3-1)(2-1)) = \chi^2(2)$ -fördelad stokastisk variabel.

$$\chi_{0.05}^2(2) = 5.99 < Q = 6.1 < \chi_{0.01}^2(2) = 9.21$$

Alltså är skillnaden signifikant på risknivån 5%, men inte på risknivån 1%. Alltså är D rätt svarsalternativ.

Uppgift 12

Vi får $\bar{x} = 43/14$ samt $\bar{y} = 572/14$. Vidare är (med beteckningar från formelsamlingen)

$$\begin{aligned} S_{xx} &= \sum_{i=1}^{14} x_i^2 - 14(\bar{x})^2 = 157.42 - 14 \cdot (43/14)^2 = 25.3486 \\ S_{xy} &= \sum_{i=1}^{14} x_i y_i - 14\bar{x}\bar{y} = 1697.80 - 14 \cdot (43/14) \cdot (572/14) = -59.0571 \\ S_{yy} &= \sum_{i=1}^{14} y_i^2 - 14(\bar{y})^2 = 23530 - 14 \cdot (572/14)^2 = 159.7143 \end{aligned}$$

Detta ger

$$\beta_{obs}^* = S_{xy}/S_{xx} = -2.3298, \quad \alpha_{obs}^* = \bar{y} - \beta_{obs}^* \bar{x} = 572/14 - (-2.3298) \cdot (43/14) = 48.0130.$$

Skattad linje är alltså $y = 48.0130 - 2.3298x$.

Vi skattar med $\alpha_{obs}^* - \beta_{obs}^* 4.3 = 48.0130 - 2.3298 \cdot 4.3 = \underline{37.9949} \approx 38$.

Del II, Lösningsförslag:**Uppgift 13**

Inför följande beteckningar: K = en enhet är korrekt; D = en enhet är defekt; \widehat{K} = en enhet klassificeras som korrekt; \widehat{D} = en enhet klassificeras som defekt. Enligt uppgiften vet vi att

$$P(D) = 0.1, \quad P(\widehat{D} | D) = 0.9, \quad P(\widehat{K} | D) = 0.1 \text{ och } P(\widehat{K} | K) = 0.85, \quad P(\widehat{D} | K) = 0.15.$$

Vi söker den betingade sannolikheten för att en enhet är defekt givet att den klassificerats som defekt, d.v.s. med beteckningarna ovan, $P(D | \widehat{D})$. För att bestämma denna sannolikhet utnyttjar vi Bayes' sats, som ger

$$P(D | \widehat{D}) = \frac{P(\widehat{D} | D) \cdot P(D)}{P(\widehat{D})}$$

Enligt lagen om total sannolikhet fås

$$P(\widehat{D}) = P(\widehat{D} | D) \cdot P(D) + P(\widehat{D} | K) \cdot P(K) = 0.9 \cdot 0.1 + 0.15 \cdot (1 - 0.1) = 0.225,$$

ty $P(K) = 1 - P(D)$. Detta ger

$$P(D | \widehat{D}) = \frac{0.9 \cdot 0.1}{0.225} = \underline{0.4}.$$

Uppgift 14

a) Alla koder väljs med lika stor sannolikhet. Vi får då att

$$\begin{aligned} P(\{\text{lyckas att knäcka koden på tre försök}\}) &= \\ 1 - P(\{\text{inte lyckas knäcka koden på tre försök}\}) &= \\ 1 - \frac{9999 \cdot 9998 \cdot 9997}{10000 \cdot 9999 \cdot 9998} &= \frac{3}{10000} = \underline{0.0003}. \end{aligned}$$

b) Låt den stokastiska variabeln X vara antalet koder som knäcks. X är då $\text{Bin}(n, p)$ -fördelad med $n = 10000$ och $p = 0.0003$. Eftersom $0.0003 \leq 0.1$ så kan vi använda approximationen $\text{Bin}(10000, 0.0003) \sim \text{Po}(3)$. Detta ger då att

$$P(X \geq 3) = 1 - P(X \leq 2) = \{\text{tabell 8}\} \approx 0.5768,$$

eller

$$\begin{aligned} P(X \geq 3) &= 1 - P(X \leq 2) = 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &\approx 1 - e^{-3} - 3e^{-3} - \frac{3^2 e^{-3}}{2!} = 1 - \frac{17e^{-3}}{2} = \underline{0.5768}. \end{aligned}$$

Om du har räknat med $p = 0.0001$ så är X approximativt $\text{Po}(1)$ -fördelad och du får på samma sätt som ovan

$$P(X \geq 3) = 1 - P(X \leq 2) = \{\text{tabell}\} \approx \underline{0.0803}.$$

Uppgift 15

Låt X vara den stokastiska variabeln "väntetid för företag A" och B "väntetid för företag B". För exponentialfördelade variabler är standardavvikelse = väntevärde, så enligt CGS (många observationer av likafördelade variabler) \bar{X} approximativt $N(\mu_A, \mu_A/\sqrt{420})$, och på samma sätt \bar{Y} approximativt $N(\mu_B, \mu_B/\sqrt{376})$. Alltså är $\bar{Y} - \bar{X}$ approximativt $N(\mu_B - \mu_A, \sqrt{\mu_A^2/420 + \mu_B^2/376})$. Eftersom μ_A och μ_B kan skattas med \bar{x} respektive \bar{y} kan vi beräkna medelfelet för $\bar{Y} - \bar{X}$ till $\sqrt{26^2/420 + 31.6^2/376} \approx 2.065$. Vi har alltså approximativt $\bar{Y} - \bar{X} \in N(\mu_B - \mu_A, 2.065)$. Ett approximativt konfidensintervall för $\mu_B - \mu_A$ är alltså

$$\mu_B - \mu_A = \bar{y} - \bar{x} \pm \lambda_{0.025} \cdot 2.065 = 5.60 \pm 4.05$$

Eftersom intervallet bara innehåller positiva tal drar vi slutsatsen att $\mu_B > \mu_A$.

Uppgift 16

Vi skall beräkna sannolikheten att intervallen $\bar{X} \pm 1.96\sigma/\sqrt{n}$ och $\bar{Y} \pm 1.96\sigma/\sqrt{n}$ inte skär varandra, där $\bar{X} = (X_1 + X_2 + \dots + X_n)/n$ och motsvarande för \bar{Y} . Eftersom X_i och Y_i är oberoende och alla $N(\mu, \sigma)$ är både \bar{X} och \bar{Y} normalfördelade $N(\mu, \sigma/\sqrt{n})$. Att intervallen inte skär varandra innebär att ett av intervallen ligger helt till vänster om det andra. Detta är ekvivalent med att $\bar{X} + 1.96\sigma/\sqrt{n} < \bar{Y} - 1.96\sigma/\sqrt{n}$ eller $\bar{Y} + 1.96\sigma/\sqrt{n} < \bar{X} - 1.96\sigma/\sqrt{n}$. Sannolikheten för dessa två händelser är på grund av symmetrin lika. Nu är händelsen $\bar{X} + 1.96\sigma/\sqrt{n} < \bar{Y} - 1.96\sigma/\sqrt{n}$ samma som $\bar{X} - \bar{Y} < -2 \cdot 1.96\sigma/\sqrt{n}$. Dessutom är $\bar{X} - \bar{Y} \in N(\mu - \mu, \sqrt{\sigma^2/n + \sigma^2/n}) = N(0, \sigma\sqrt{2/n})$. Det ger oss

$$P(\bar{X} - \bar{Y} < -2 \cdot 1.96\sigma/\sqrt{n}) = \Phi\left(\frac{-2 \cdot 1.96\sigma/\sqrt{n}}{\sigma\sqrt{2/n}}\right) = \Phi(-1.96\sqrt{2}) = \Phi(-2.7719) = 1 - \Phi(2.7719) = 0.0028.$$

Sannolikheten att intervallen inte skär varandra blir således $2 \cdot 0.0028 = \underline{0.0056}$.