Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF1923 OCH SF1924 SANNOLIKHETSTEORI OCH STATISTIK, SAMT TENTAMEN I SF1935 SANNOLIKHETSTEORI OCH STATISTIK MED TILLÄMPNING INOM MASKININLÄRNING. ONSDAG 13 AUGUSTI 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 ordinarietentamen i maj 2025 och vid omtentamen i augusti 2025. Gränsen för godkänt är 9 poäng på del I. Möjlighet att komplettera ges för tentander med 8 poäng på del I.
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 ordinarietentamen i maj 2025 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

I ett stort bostadsområde är $\frac{1}{6}$ av lägenheterna ettor, $\frac{1}{3}$ tvåor, $\frac{2}{5}$ treor och $\frac{1}{10}$ fyror. Under en 10-årsperiod anses risken för vattenskada för en etta vara 10% , för en tvåa 5%, för en trea 8 % och för en fyra 15%. Anta att en vattenskada inträffat. Vad är då sannolikheten att det är en tvåa som har drabbats?

A: 0.010

B: 0.069

C: 0.095

D: 0.207

Uppgift 2

En stokastisk variabel X har täthetsfunktionen

$$f_X(x) = \begin{cases} \frac{1}{4}, & 0 < x < 2 \\ \frac{1}{8}, & 4 < x < 8 \\ 0, & \text{annars.} \end{cases}$$

Bestäm standardavvikelsen $D(X)$.

A: 1.29

B: 1.73

C: 2.66

D: 3.46

Uppgift 3

De stokastiska variablerna X och Y har den simultana sannolikhetsfunktionen

$p_{X,Y}(x,y)$	$X = 1$	$X = 2$	$X = 3$
$Y = 0$	0.1	0.2	0.2
$Y = 1$	0.1	0.1	0.3

Således antar X värdena 1, 2 och 3, medan Y antar värdena 0 och 1.
Beräkna $V(X + Y)$.

A: 0.86

B: 0.96

C: 2.40

D: 4.20

Uppgift 4

En elektronisk komponents livslängd är exponentialfördelad med väntevärdet fyra timmar. Efter tre timmars gångtid fungerar fortfarande komponenten. Beräkna sannolikheten att komponenten fungerar åtminstone två timmar till.

A: $3.35 \cdot 10^{-4}$

B: 0.18

C: 0.29

D: 0.61

Uppgift 5

De stokastiska variablerna X och Y är oberoende, X är $\mathrm{Bin}(10, 0.1)$ och Y är $\mathrm{Bin}(5, 0.1)$. Sätt $Z = X + 2Y$. Beräkna $P(Z = 2)$.

A: $1.50 \cdot 10^{-3}$

B: 0.0254

C: 0.229

D: 0.285

Uppgift 6

Vid en statistisk kvalitetskontroll tas 250 enheter ut för kontroll. Antalet fel på en enhet är poissonfördelat med väntevärde μ och antalet fel på olika enheter är oberoende av varandra. Antag att $\mu = 0.3$.
Sammanlagda antalet fel på de 250 enheterna beräknas (s.k. felantalskontroll). Partiet accepteras om högst 90 fel finns. Vad är den approximativa sannolikheten att partiet accepteras?

A: 0.58

B: 0.69

C: 0.80

D: 0.96

Uppgift 7

Antag att en preliminär undersökning av vad man tycker om att få installera solcellspaneler hos sig har gjorts och att av $n = 1000$ tillfrågade svarade $x = 600$ att de vill installera solcellspaneler. Man skattar därför andelen p som vill göra det med $p^*_{obs} = 600/1000 = 0.6$.

Bestäm medelfelet för p^*, dvs bestäm $d(p^*)$.

A: 0.4

B: 0.0245

C: 0.0155

D: 0.49

Uppgift 8

Kundserviceavdelningen hos ett visst företag överväger att anställa fler personer för att svara i telefon när deras kunder ringer i olika ärenden. Det är därför av intresse att modellera antalet inkommande samtal under en arbetsdag, vilket antas följa en Poissonfördelning med okänd parameter μ. För att skatta μ har avdelningen räknat det totala antalet samtal under 25 dagar och ser då att 307 olika samtal har registrerats under den perioden. Bestäm ett approximativt tvåsidigt konfidensintervall med konfidensgrad 95% för μ. Vad är den undre gränsen för detta intervall?

A: 10.91

B: 11.13

C: 5.41

D: 6.52

Uppgift 9

En förskare har gjort tio försök som anses vara oberoende av varandra där sannolikheten för lyckat försök är p. Låt X stå för antalet lyckade försök. Förskaren önskar pröva $H_0 : p = 1/2$ mot $H_1 : p = 4/5$ och förkastar H_0 till förmån för H_1 om man får fler än sju lyckade försök. Resultatet av de tio försöken var nio lyckade försök. Beräkna P-värdet!

A: 0.001

B: 0.0107

C: 0.0547

D: 0.6778

Uppgift 10

Biverkningseffekter av två typer av medicin, benämnda Typ 1 och Typ 2, har studerats i en grupp om 570 personer. Resultatet sammanfattas i tabellen nedan.

	Medicin av Typ 1	Placebo	Medicin av Typ 2
Antal som upplevde biverkningar	100	30	20
Antal som inte upplevde biverkningar	300	50	70

För att testa nollhypotesen H_0 : Det finns ingen skillnad mellan grupperna, beräknar man teststorheten Q och får $Q = 6.296$. Vilken slutsats kan man dra då man fått denna teststorhet?

A: H_0 kan förkastas på risknivån 1%, men inte på risknivån 5%

B: H_0 kan förkastas på både risknivån 1% och risknivån 5%

C: H_0 kan varken förkastas på risknivån 1% eller risknivån 5%

D: H_0 kan förkastas på risknivån 5%, men inte på risknivån 1%

Uppgift 11

Antag att $X_1, \ldots, X_n$ utgör ett stickprov på $N(\mu, \sigma)$, där σ är okänd. Ebba önskar testa nollhypotesen $H_0 : \mu = 2$ mot $H_1 : \mu > 2$ med hjälp av ett lämpligt konfidensintervall för μ.
Vilket av nedanstående konfidensintervall för μ skall väljas för att testets signifikansnivå skall bli α?

A: $I_\mu = \left(-\infty,\ \overline{x} + \frac{s}{\sqrt{n}} \cdot t_\alpha(n-1)\right)$

B: $I_\mu = \left(\overline{x} - \frac{s}{\sqrt{n}} \cdot t_\alpha(n-1), \infty\right)$

C: $I_\mu = \left(-\infty,\ \overline{x} + \frac{s}{\sqrt{n}} \cdot t_{\alpha/2}(n-1)\right)$

D: $I_\mu = \left(\overline{x} - \frac{s}{\sqrt{n}} \cdot t_{\alpha/2}(n-1), \infty\right)$

Uppgift 12

I ett avancerat växthus utförs ett experiment för att avgöra om mängden belysning påverkar hur mycket jordgubbar växer. Belysningen mäts med hjälp av ett belysningsindex och jordgubbarnas vikt mäts i gram. De första fyra erhållna observationerna följer nedan

Belysningsindex	5	10	10	15
Jordgubbsvikt (g)	20	26	34	40

Det är rimligt att tro att det föreligger ett linjärt samband mellan variablerna. Utifrån datamaterialet ovan skattas en linjär regressionsmodell

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

där y_i =jordgubbsvikt (g) beror av x_i = belysningsindex och ε_i betecknar slumpmässiga fel, $i = 1, \ldots, 4$. Minsta-kvadrat skattningarna av regressionskoefficienterna α och β blev $\alpha^*_{obs} = 10$ respektive $\beta^*_{obs} = 2$.
Vilket av de fyra svarsalternativen nedan motsvarar ett 90 % tvåsidigt konfidensintervall för den effekt som belysningsindex har på jordgubbsvikten, dvs $I_\beta = I_\beta(0.9)$, om man vet att effekten i fråga är signifikant på 10 % nivå.

A: [−0.35, 3.35]

B: [0.35, 3.65]

C: [−0.43, 4.43]

D: [0.35, 4.65]

Var god vänd!

Del II

Uppgift 13

I ett stort bostadsområde är $\frac{1}{6}$ av lägenheterna ettor, $\frac{1}{3}$ tvåor, $\frac{2}{5}$ treor och $\frac{1}{10}$ fyror. Bestäm sannolikheten (approximativt) att 60 slumpvis utvalda lägenheter tillsammans har minst 135 rum. (10 p)

Uppgift 14

Ett litet delsystem i ett tekniskt system består av tre parallellkopplade identiska komponenter. Komponenternas livslängder kan betraktas som oberoende stokastiska variabler X_1, X_2 resp. X_3, var och en med täthetsfunktionen

$$f(x) = \begin{cases} \frac{x}{2}e^{-x^2/4}, & \text{för } x > 0; \\ 0, & \text{för } x \leq 0. \end{cases}$$

Låt Y beteckna den första tidpunkt då någon av komponenterna slutar att fungera, och Z delsystemets livslängd, d.v.s. den tidpunkt då alla komponenterna har slutat att fungera.

a) Beräkna $P(Z \leq 2)$. (5 p)

b) Beräkna $P(Z \leq 4|Y > 2)$. (5 p)

Uppgift 15

Ett visst mätinstrument ger slumpmässiga fel. Mätfelen vid 16 upprepade mätningar kan betrkatas som oberoende obervationer $x_1, x_2, \ldots, x_{16}$ från en normalfördelning $N(\mu, \sigma)$, där μ är okänd och $\sigma = 1$ (enhet: mm). Man vill veta om $\mu = 0$, eller om instrumentet i genomsnitt ger ett mätfel som är större än noll. Man vill därför ha ett test av $H_0 : \mu = 0$ mot $H_1 : \mu > 0$.

(a) Anna använder medelvärdet $\bar{x}$ som testvariabel och förkastar H_0 om $\bar{x} > c$, där c är en konstant. Vilket värde på c gör att testet får 1 % signifikansnivå ? (5 p)

(b) Stefan avser att testa samma hypotes med en annan signifikansnivå. Hans beslutsregel är: Förkasta H_0 om $\bar{x} > 0.683$. Beräkna styrkan för Stefans test i punkten $\mu = 0.9$. (5 p)

Uppgift 16

Låt $x_1, x_2, \ldots, x_n$ vara observationer av de oberoende stokastiska variablerna $X_1, X_2, \ldots, X_n$ vilka alla är likformig fördelade $U(0, b)$, där b är okänd. Man kan visa att den korrigerade ML-skattningen av b, som är en väntevärdesriktig skattning, ges av

$$b^*_{obs} = c(n) \max_{1 \leq k \leq n} x_k \, ,$$

där $c(n) = (n + 1)/n$. Härled variansen $V(b^*)$ och bestäm sedan medelfelet för b^* då $x_1 = 2.0$, $x_2 = 9.5$ och $x_3 = 4.4$. (10 p)

Lycka till!

Avd. Matematisk statistik

KTH Matematik

LÖSNINGSFÖRSLAG
TENTAMEN I SF1923 OCH SF1924 SANNOLIKHETSTEORI OCH STATISTIK, SAMT TENTAMEN I SF1935 SANNOLIKHETSTEORI OCH STATISTIK MED TILLÄMPNING INOM MASKININLÄRNING
ONSDAG 13 AUGUSTI 2025 KL 8.00–13.00.

Del I, Svar.

1.D

2.C

3.B

4. D

5. C

6. D

7. C

8. A

9. B

10. D

11. B

12. B

Del I, Lösningsförslag.

Uppgift 1

Här har vi Bayes sats.
Låt A, B, C, D beteckna händelserna att en lägenhet har ett, två, tre resp. fyra rum.
Låt V beteckna händelsen att en vattenskada inträffat. Då blir

$$P(B|V) = \frac{P(B \cap V)}{P(V)} = \frac{P(V|B)P(B)}{P(V|A)P(A) + P(V|B)P(B) + P(V|C)P(C) + P(V|D)P(D)} =$$

$$= \frac{0.05 \cdot \frac{1}{3}}{0.10 \cdot \frac{1}{6} + 0.05 \cdot \frac{1}{3} + 0.08 \cdot \frac{2}{5} + 0.15 \cdot \frac{1}{10}} = \frac{50}{241} = 0.207$$

Svar: <u>Sannolikheten är 20.7% att det är en tvåa som har drabbats.</u>

Uppgift 2

$$E(X) = \int_0^2 \frac{x}{4} dx + \int_4^8 \frac{x}{8} dx = \left[\frac{x^2}{8}\right]_0^2 + \left[\frac{x^2}{16}\right]_4^8 = \frac{1}{2} + 4 - 1 = \frac{7}{2}$$

$$E(X^2) = \int_0^2 \frac{x^2}{4} dx + \int_4^8 \frac{x^2}{8} dx = \left[\frac{x^3}{12}\right]_0^2 + \left[\frac{x^3}{24}\right]_4^8 = \frac{8}{12} + \frac{512}{24} - \frac{64}{24}$$

$$V(X) = E(X^2) - E^2(X) = \frac{16}{24} + \frac{512}{24} - \frac{64}{24} - \frac{49}{4} = \frac{170}{24}$$

$$D(X) = \sqrt{\frac{170}{24}} = 2.66$$

Uppgift 3

Sätt $Z = X + Y$. Då får vi följande sannolikhetsfunktion för Z.

$$p_Z(z) = \begin{cases} 0.1, & z = 1 \\ 0.3, & z = 2 \\ 0.3, & z = 3 \\ 0.3, & z = 4 \end{cases}$$

$$E[Z] = \sum_{k=1}^{4} k \cdot p_Z(k) = 0.1 + 0.6 + 0.9 + 1.2 = 2.8$$

$$E[Z^2] = \sum_{k=1}^{4} k^2 \cdot p_Z(k) = 0.1 + 1.2 + 2.7 + 4.8 = 8.8$$

$$V(X) = E(X^2) - E^2(X) = 8.8 - 2.8^2 = 0.96$$

Uppgift 4

För komponentens livslängd X gäller $P(X \leq x) = 1 - e^{-x/4}$, $x > 0$. Den sökta sannolikheten blir

$$P(X > 3 + 2 \,|\, X > 3) = P(X > 5)/P(X > 3) = e^{-5/4}/e^{-3/4} = e^{-1/2} = 0.61.$$

Uppgift 5

$$P(Z = 2) = P(X = 2 \cap Y = 0) + P(X = 0 \cap Y = 1)$$
$$= \binom{10}{2} 0.1^2 \cdot 0.9^8 \binom{5}{0} 0.1^0 \cdot 0.9^5 + \binom{10}{0} 0.1^0 \cdot 0.9^{10} \binom{5}{1} 0.1 \cdot 0.9^4 = 0.9^{14} = 0.23.$$

Uppgift 6

Sätt X=totala antalet fel i urvalet. X är Po($250 \cdot 0.3$)=Po(75) eftersom summan av oberoende poissonfördelade stokastiska variabler själv är poissonfördelad. Men Po(75)≈N($75, \sqrt{75}$) ty $75 \geq 15$ varför vi erhåller

$$P(X \leq 90) \approx \Phi\left(\frac{90 - 75}{\sqrt{75}}\right) \approx \Phi(1.732) \approx \underline{0.958} \approx \underline{0.96}$$

Uppgift 7

Låt stokastisk variabel $X \in Bin(n, p)$, där p är okänd. Vi vet att av $n = 1000$ tillfrågade svarade $x = 600$ att dom vill installera solcellpaneler, dvs x är en observation på s.v. X. Man skattar p med $p^* = X/n$. Därför har vi:

$$V(p^*) = V\left(\frac{X}{n}\right) = \frac{1}{n^2}\big(V(X)\big) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}\,.$$

Medelfelet blir

$$d(p*) = \big(D(p^*)\big)_{obs} = \left(\sqrt{\frac{p(1-p)}{n}}\right)_{obs} = \sqrt{\frac{p_{obs}(1-p_{obs})}{n}} = |p_{obs} = 600/1000 = 0.6,\ n = 1000| =$$

$$= \sqrt{\frac{0.6(1-0.6)}{1000}} = 0.0155$$

Svaret är C.

Uppgift 8

Vi har:

$$I_\mu = I_\mu(0.95) = \left(307/25 \pm 1.96\sqrt{\frac{307/25}{25}}\right),$$

där den undre gränsen är lika med

$$307/25 - 1.96\sqrt{\frac{307/25}{25}} \approx 10.91$$

Svaret är A.

Uppgift 9

Vi har:

$$P\text{-värdet} = P\left(X \geq 9 | H_0 \text{ sann}\right) = 1 - P\left(X \leq 8 | H_0 \text{ sann}\right) = \lceil \text{Tabell 6 där } n = 10, p = \frac{1}{2} \rceil =$$

$$= 1 - 0.98926 = 0.0107$$

Svaret är B.

Uppgift 10

Vi har $(r-1)(c-1) = (2-1)(3-1) = 2$ frihetsgarder. Eftersom teststorheten $Q = 6.296 > 5.99 = \chi^2_{0.05}(2)$ (se Tabell 4, för $\chi^2(2)$-fördelning), så därför H_0 kan förkastas på risknivån 5%. Dock inte på nivån 1% (eftersom $Q = 6.296 < 9.21 = \chi^2_{0.01}(2)$).

Svaret därför är D.

Uppgift 11

Eftersom signifikansnivån skall vara α så kan varken alternativ C eller alternativ D vara rätt svar. P.g.a mothypotesen H_1 måste då det rätta svaret vara alternativ B.

Uppgift 12

1) Konfidensintervallet I_β måste innehålla $\beta^*_{obs} = 2$, eftersom intervallet $I_\beta = I_\beta(0.9)$ är symmetriskt runt $\beta^*_{obs} = 2$.
2) Konfidensintervallet I_β ska *inte* innehålla noll, eftersom effekten som belysningsindex har på jordgubbsvikten är signifikant.

Det finns bara ett intervall (nämligen, i svarsalternativ B) som uppfyller de två kraven (dvs kraven 1) och 2)) ovan.

Svaret är B.

Del II, Lösningsförslag.

Uppgift 13

Låt X_i vara den stokastiska variabeln *antal rum* för en på måfå observerad lägenhet. Vi får då

$$\mu = E[X_i] = \sum_{k=1}^{4} k \cdot p_{X_i}(k) = 1\frac{1}{6} + 2\frac{1}{3} + 3\frac{2}{5} + 4\frac{1}{10} = \frac{146}{60} \approx 2.433$$

och

$$\sigma = D(X_i) = \sqrt{V(X_i)}$$

där

$$V(X_i) = E(X_i^2) - E^2(X_i) = E(X_i^2) - (\frac{146}{60})^2$$

och

$$E(X_i^2) = \sum_{k=1}^{4} k^2 \cdot p_{X_i}(k) = 1^2\frac{1}{6} + 2^2\frac{1}{3} + 3^2\frac{2}{5} + 4^2\frac{1}{10} = \frac{402}{60}$$

Detta medför att

$$V(X_i) = \frac{402}{60} - (\frac{146}{60})^2 = \frac{2804}{3600} \approx 0.779$$

Vilket ger att

$$\sigma = D(X_i) = \sqrt{\frac{2804}{3600}} \approx 0.883$$

Låt nu

$$Y = \sum_{i=1}^{60} X_i$$

Vi antar nu att X_i:na är många och oberoende. Eftersom de dessutom är likafördelade gäller då enligt Centrala Gränsvärdessatsen C.G.S. att

$$Y \sim N(60\mu, \sigma\sqrt{60}) = N(60\frac{146}{60}, \sqrt{\frac{2804}{3600}}\sqrt{60}) = N(146, \sqrt{\frac{2804}{60}})$$

$$P(Y \geq 135) = P(\frac{Y-146}{\sqrt{\frac{2804}{60}}} \geq \frac{135-146}{\sqrt{\frac{2804}{60}}}) = [Z = \frac{Y-146}{\sqrt{\frac{2804}{60}}} \sim N(0,1)] = P(Z \geq \frac{135-146}{\sqrt{\frac{2804}{60}}}) =$$

$$= 1 - P(Z \leq \frac{135 - 146}{\sqrt{\frac{2804}{60}}}) = 1 - \Phi(\frac{-11}{\sqrt{\frac{2804}{60}}}) = 1 - [1 - \Phi(\frac{11}{\sqrt{\frac{2804}{60}}})] \approx \Phi(1.61) \approx 0.946$$

Svar: <u>Sannolikheten att antalet rum är minst 135 är ungefär 94.6%</u>

Uppgift 14

Vi observerar att $Y = \min\{X_1, X_2, X_3\}$ och $Z = \max\{X_1, X_2, X_3\}$.

a)

$$P(Z \leq 2) = P(X_1 \leq 2 \cap X_2 \leq 2 \cap X_3 \leq 2) = P(X_1 \leq 2)^3$$

$$= \Big(\int_0^2 \frac{x}{2} e^{-x^2/4} dx\Big)^3 = (1 - e^{-1})^3 = \underline{0.2526}.$$

b)

$$P(Z \leq 4 \mid Y > 2) = \frac{P(Z \leq 4 \cap Y > 2)}{P(Y > 2)} = \frac{P(2 < X_1 \leq 4 \cap 2 < X_2 \leq 4 \cap 2 < X_3 \leq 4)}{P(X_1 > 2 \cap X_2 > 2 \cap X_3 > 2)}$$

$$= \frac{P(2 < X_1 \leq 4)^3}{P(X_1 > 2)^3} = \frac{\big(\int_2^4 \frac{x}{2} e^{-x^2/4} dx\big)^3}{\big(\int_2^\infty \frac{x}{2} e^{-x^2/4} dx\big)^3} = \frac{(e^{-1} - e^{-4})^3}{e^{-3}} = (1 - e^{-3})^3 = \underline{0.8580}.$$

Uppgift 15

(a) Eftersom $H_1 : \mu > 0$, förakstar man $H_0 : \mu = 0$ på signifikansnivån 1% om

$$\frac{\bar{x} - 0}{\sigma/\sqrt{n}} > \lambda_{0.01}\,,$$

där $n = 16$, $\sigma = 1$ och $\lambda_{0.01} = 2.3263$ enl Tabell 2. Då har vi efter omskrivningen:

$$\bar{x} > \lambda_{0.01} \frac{\sigma}{\sqrt{n}},$$

och därför

$$c = \lambda_{0.01} \frac{\sigma}{\sqrt{n}} = \frac{2.3263}{\sqrt{16}} = 0.581575$$

Svar: $c = 0.5816$

(b) Vi har:

$$h(0.9) = P(\text{förkasta } H_0 | H_1 \text{ är sann}) = P(\bar{X} > 0.683 | \mu = 0.9) =$$

$$= P(\bar{X} > 0.683 | \bar{X} \in N(0.9, \sigma/\sqrt{n})) = 1 - \Phi\Big(\frac{0.683 - 0.9}{1/\sqrt{16}}\Big) =$$

$$= 1 - \Phi\big(-0.868\big) = \Phi\big(0.868\big) \approx \Phi\big(0.87\big) = 0.8078\,,$$

enligt Tabell 1.

Svar: $h(0.9) = 0.8078$

Uppgift 16

Vi vet att

$$b^*_{obs} = c(n) \max_{1\le k\le n} x_k \,,$$

där $c(n) = (n+1)/n$, är en väntevärdesriktig skattning av b. DVS.,

$$E(b^*) = c(n)E\big(\max_{1\le k\le n} X_k\big) = b\,.$$

Följaktligen:

$$V(b^*) = E\big((b^*)^2\big) - \big(E(b^*)\big)^2 = \frac{(n+1)^2}{n^2}\, E\Big(\max_{1\le k\le n} X_k\Big)^2 - b^2\,.$$

Vidare, fördelningsfunktionen för stokastisk variabel $\max_{1\le k\le n} X_k$ blir:

$$F_{\max_{1\le k\le n} X_k}(y) = P\big(\max_{1\le k\le n} X_k \le y\big) = P\big(\text{ alla } X_k \le y\big) = |\text{oberoendet}| =$$

$$= \prod_{k=1}^{n} P(X_k \le y) = |\text{lika fördelning}| = \Big(F_{X_1}(y)\Big)^n,$$

och därmed blir täthetsfunktionen för s.v. $\max_{1\le k\le n} X_k$ lika med:

$$f_{\max_{1\le k\le n} X_k}(y) = \frac{dF_{\max_{1\le k\le n} X_k}(y)}{dy} = n\Big(F_{X_1}(y)\Big)^{n-1} f_{X_1}(y) = \frac{n}{b}\left(\frac{y}{b}\right)^{n-1} = n\,\frac{y^{n-1}}{b^n}\,,$$

där $0 \le y \le b$, ty $X_1 \in U(0,b)$. Därför har vi:

$$E\Big(\max_{1\le k\le n} X_k\Big)^2 = \int_0^b y^2\, n\,\frac{y^{n-1}}{b^n}\, dy = |\text{enkla beräkningar}| = \frac{n}{n+2}\, b^2\,,$$

och därmed,

$$V(b^*) = \frac{(n+1)^2}{n^2}\, E\Big(\max_{1\le k\le n} X_k\Big)^2 - b^2 = b^2\left(\frac{(n+1)^2 n}{n^2(n+2)} - 1\right) = ... = \frac{1}{n(n+2)}\, b^2\,.$$

Då $x_1 = 2.0$, $x_2 = 9.5$ och $x_3 = 4.4$, blir medelfelet $d(b^*)$ för b^* lika med:

$$d(b^*) = \Big(\sqrt{V(b^*)}\Big)^*_{obs} = b^*_{obs}\sqrt{\frac{1}{n(n+2)}} = \big((n+1)/n\big)\,\Big(\max_{1\le k\le n} x_k\Big)\sqrt{\frac{1}{n(n+2)}} =$$

$$= |n=3| = (4/3)\,(9.5)\,\sqrt{1/15} = 3.27\,.$$

Och själva korrigerade ML-skattningen b^*_{obs} av b är:

$$b^*_{obs} = \big((n+1)/n\big)\,\max_{1\le k\le n} x_k = (4/3)(9.5) = 12.67\,.$$