KTH Matematik

Avd. Matematisk statistik

TENTAMEN I SF1920/SF1921 SANNOLIKHETSTEORI OCH STATISTIK,
ONSDAG 5 JUNI 2019 KL 14.00–19.00.

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. På denna del skall endast svar anges, antingen i form av ett numeriskt värde med tre värdesiffrors noggrannhet eller i form av val av ett av de möjliga svarsalternativen. Studenter som är godkända på kontrollskrivningen behöver ej besvara uppgift 1-3, utan får tillgodoräkna sig dessa tre uppgifter. Gränsen för godkänt är preliminärt 9 poäng. Möjlighet att komplettera ges för tentander med, preliminärt, 8 poäng. Tid och plats för komplettering kommer att anges på kursens hemsida. Del II består av uppgifterna 13-16 och varje korrekt lösning ger 10 poäng. Del II rättas bara för studenter som är godkända på del I och poäng på del II krävs för högre betyg än E. På denna del skall resonemang och uträkningar skall vara så utförliga och väl motiverade att de är lätta att följa. Införda beteckningar skall förklaras och definieras och numeriska svar skall anges med minst två värdesiffrors noggrannhet. Studenter som är godkända på datorlaborationen får 4 bonuspoäng på del II på ordinarie tentamenstillfället och det första omtentamenstillfället.

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 händelser sådana att $P(B) = 1/5$, $P(A|B) = 1/6$ och $P(A|B^*) = 1/7$.
Beräkna $P(B^*|A)$.

A: 0.226

B: 0.148

C: 0.774

D: 0.033

Uppgift 2

En stokastisk variabel X har täthetsfunktionen

$$f_X(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 \le x \le 1 \\ \frac{1}{3}, & 1 < x \le 3 \\ 0, & x > 3 \end{cases}$$

Bestäm väntevärdet för X^3, dvs $E(X^3)$.

Uppgift 3

De tre stokastiska variablerna X, Y och Z uppfyller $V(X) = V(Y) = V(Z) = 1$, samt $C(X, Y) = C(X, Z) = C(Y, Z) = 1/2$. Låt $W = X + Y + Z$. Beräkna $D(W + 3)$.

A: 2.45

B: 3.00

C: 2.12

D: 1.73

Uppgift 4

Låt X och Y vara två oberoende stokastiska variabler sådana att $X \in \text{Bin}(5, 0.2)$ och $Y \in \text{Po}(1)$. Beräkna $P(X + Y = 2)$.

A: 0.678

B: 0.136

C: 0.286

D: 0.151

Uppgift 5

Antag att $X \in N(6, 2)$. Beräkna sannolikheten $P(4 \le X \le 8)$.

A: 1.000

B: 0.383

C: 1.683

D: 0.683

Uppgift 6

Låt $X \in \text{Exp}(1/4)$, d v s intensiteten är en fjärdedel. Bestäm sannolikheten att X antar ett värde större än 5, givet att värdet är större än 2.

A: 0.528

B: 0.287

C: 0.320

D: 0.472

Uppgift 7

Låt $\bar{x} = 137.0$, $\bar{y} = 208.0$, $s_x^2 = 92.5$, $s_y^2 = 163.0$, $n_x = 500$ och $n_y = 300$ vara givet och antag att X_i:na alla har samma okända fördelning, och att Y_j:na alla har samma okända fördelning (kanske dock ej samma som X_i:na). Antag vidare att alla dessa stokastiska variabler är oberoende.
Ange undre gränsen för det approximativt 99%-iga tvåsidiga konfidensintervallet $I_{\mu_x - \mu_y}$.

A: -97.48

B: -94.91

C: -73.20

D: -72.99

Uppgift 8

Antag att $X_1, \ldots, X_n$ utgör ett stickprov på $N(\mu, \sigma)$, där σ är känd. Evert 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 signaifikansnivå skall bli α?

A: $I_\mu = \left(-\infty,\ \overline{x} + \dfrac{\sigma}{\sqrt{n}} \cdot \lambda_\alpha\right)$

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

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

D: $I_\mu = \left(\overline{x} - \dfrac{\sigma}{\sqrt{n}} \cdot \lambda_\alpha, \infty\right)$

Uppgift 9

Antag att att vi har två oberoende observationer x_1 och x_2 av $X_i \in Bin(n_i, p)$, $i = 1, 2$. Ange ML-skattningen av p.

A: $\frac{1}{2}\left(\frac{x_1}{n_1} + \frac{x_2}{n_2}\right)$

B: $\frac{x_1 + x_2}{n_1 + n_2}$

C: $\frac{n_1 x_1 + n_2 x_2}{n_1^2 + n_2^2}$

D: $\frac{n_2^2 x_1 + n_1^2 x_2}{n_1 + n_2}$

Uppgift 10

Antag att vi har N stokastiska variabler X_i som är oberoende och tillhör Γ -fördelningen $\Gamma(k, \Theta)$.

$$\text{Då är} \quad E(X_i) = \frac{k}{\Theta} \quad \text{och} \quad V(X_i) = \frac{k}{\Theta^2}$$

$$\text{Maximum-likelihood-skattningen av} \quad \Theta \quad \text{är} \quad \Theta^*_{obs_{ML}} = \frac{\Sigma x_i}{k \cdot N}$$

Vad är variansen av denna skattning?

A: $\frac{1}{k \cdot N \cdot \Theta^2}$

B: $\frac{1}{N^2 \cdot \Theta^2}$

C: $\frac{1}{N \cdot \Theta^2}$

D: $\frac{1}{k^2 \cdot N^2 \cdot \Theta^2}$

Uppgift 11

Vad menas med att en skattning av en parameter är väntevärdesriktig?

A: Skattningen sammanfaller med stickprovsmedelvärdet.

B: Skattningen är ett viktat medelvärde av observationerna.

C: Skattningen ger alltid rätt parametervärde.

D: Väntevärdet av stickprovsvariabeln är lika med det rätta parametervärdet.

Uppgift 12

Man misstänkte att ett roulettebord på ett kasino var manipulerat och genomförde ett test med 8000 försök. Om rouletten är korrekt skall röd, svart och grön (nollan) komma upp i proportionerna 18:18:1. Testresultatet gav röd: 3771, svart: 4002, grön: 227.
För att testa nollhypotesen H_0 :

$$P(\text{röd}) = \frac{18}{37}, P(\text{svart}) = \frac{18}{37}, \quad P(\text{grön}) = \frac{1}{37},$$

beräknar man teststorheten Q och får $Q = 7.41$.
Bestäm testets P-värde.

A: 1%

B: 2.46%

C: 97.5%

D: 5%

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?

Uppgift 14

I ett planerat bostadsområde med 80 lägenheter bygger man 80 parkeringsplatser. Hur stor är den approximativa sannolikheten att antalet parkeringsplatser skall räcka till, om vi antar att sannolikheten att en slumpvald lägenhet i denna typ av bostadsområde behöver 0 parkeringsplatser, 1 parkeringsplats respektive 2 parkeringsplatser är 0.4, 0.4 respektive 0.2?

Uppgift 15

Vid en undersökning av om hållfastheten hos en viss typ av cement verkar bero av härdningstiden bestämdes hållfastheten hos 14 olika provkroppar, som hade härdats under 2 respektive 7 dagar. Man fick följande resultat (i kp/m^2)

Härdningstid	Hållfasthet						
2 dagar	21.8	21.7	20.0	22.5	22.0	22.1	21.9
7 dagar	32.4	31.8	34.5	33.9	34.4	34.2	34.9

Hållfastheten vid de båda härdningstiderna kan antas vara normalfördelad med samma standardavvikelse σ, och vi kan anta oberoende mellan samtliga observationer. Testa om det verkar spela någon roll för hållfastheten hos denna typ av cement om vi härdar den 2 dagar eller 7 dagar. Använd signifikansnivån 10%. Ange tydligt vilka de uppställda hypoteserna är. Din slutsats skall tydligt motiveras. (10 p)

Uppgift 16

För att bestämma längden av en cirkels diameter d har man gjort n mätningar av diametern och m mätningar av cirkelns omkrets. Mätningarna kan uppfattas som utfall av oberoende normalfördelade stokastiska variabler vilkas väntevärden överensstämmer med de sanna längderna och vilkas standardavvikelse är σ, där σ är ett känt tal. Detta svarar emot att mätutrustningen inte har något systematiskt fel och att den är beprövad så att man vet standardavvikelsen.

Skatta cirkelns diameter d med MK-metoden utgående från

1. de n mätningarna av cirkelns diameter,
2. de m mätningarna av cirkelns omkrets,
3. alla $n + m$ mätningar.

Man kan visa att om θ^*_{obs} och $\widehat{\theta}_{obs}$ är två oberoende och väntevärdesriktiga punktskattningar av θ med de kända varianserna σ_1^2 respektive σ_2^2 så fås den mest effektiva skattningen på formen

$$\widetilde{\theta}_{obs} = a\theta^*_{obs} + (1-a)\widehat{\theta}_{obs} \tag{1}$$

då

$$a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}.$$

Använd ovanstående för att visa att MK-skattningen av d baserat på alla $n + m$ mätningar är den effektivaste skattningen man kan få då man kombinerar ihop skattningen baserat på de n mätningarna av cirkelns diameter och de m mätningarna av cirkelns omkrets enligt formel (1). Du behöver inte kontrollera väntevärdesriktigheten. (10 p)

Lycka till!

Avd. Matematisk statistik

KTH Matematik

LÖSNINGSFÖRSLAG TENTAMEN I SF1920/SF1921 SANNOLIKHETSTEORI OCH STATISTIK, ONSDAG 5 JUNI 2019 KL 14.00–19.00.

Del I

Uppgift 1

Eftersom B och B^* utgör en partition av Ω kan vi använda lagen om total sannolikhet på A

$$P(A) = P(B)\,P(A \mid B) + P(B^*)\,P(A \mid B^*) = \frac{1}{5} \times \frac{1}{6} + \frac{4}{5} \times \frac{1}{7} = \frac{155}{30 \times 35}$$

Enligt definitionen av betingad sannolikhet har vi

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B)\,P(A \mid B)}{P(A)} = \frac{\frac{1}{5} \times \frac{1}{6}}{\frac{155}{30 \times 35}} = \frac{35}{155} = 0.226$$

Till slut använder vi lagen om komplementhändelsen på den betingade sannolikheten för B^* givet A.

$$P(B^*|A) = 1 - P(B|A) = 1 - \frac{35}{155} = \frac{120}{155} = 0.774$$

Uppgift 2

Väntevärdet blir

$$\begin{aligned}
E\left(X^3\right) &= \int_{-\infty}^{\infty} x^3 \cdot f_X(x)\,dx \\
&= \int_0^1 x^3 \cdot x^2 dx + \int_1^3 x^3 \cdot \frac{1}{3} dx \\
&= \left[\frac{x^6}{6}\right]_0^1 + \left[\frac{x^4}{12}\right]_1^3 \\
&= \frac{1}{6} + \left[\frac{3^4}{12} - \frac{1}{2}\right] \\
&= = \frac{1}{6} + \frac{80}{12} = 41/6 \approx 6.83
\end{aligned}$$

Uppgift 3

Vi har att $D(W+3) = D(W)$. Vi beräknar först variansen för W och tar kvadratroten sist.

$$\begin{aligned}
V(X+Y+Z) &= V(X) + V(X) + V(X) \\
&+ 2C(X,Y) + 2C(X,Z) + 2C(Y,Z) \\
&= 3 \times 1 + 3 \times 2 \times \frac{1}{2} \\
&= 6
\end{aligned}$$

Alltså blir $D(W) = \sqrt{6} = 2.45$

Uppgift 4

$P(X+Y=2) = [\text{ober}] = p_X(0) \cdot p_Y(2) + p_X(1) \cdot p_Y(1) + p_X(2) \cdot p_Y(0) =$
$= 0.32768 \cdot 0.18394 + 0.40960 \cdot 0.36788 + 0.20480 \cdot 0.36788 = 0.286$

Uppgift 5

$$\begin{aligned}
P(4 \leq X \leq 8) &= P\left(\frac{4-6}{2} \leq \frac{X-6}{2} \leq \frac{8-6}{2}\right) \\
&= P\left(-1 \leq \frac{X-6}{2} \leq 1\right) \\
&= \Phi(1) - \Phi(-1) \\
&= \Phi(1) - (1 - \Phi(1)) \\
&= 2 \times \Phi(1) - 1 \\
&= 2 \times 0.8413 - 1 \\
&= 0.6826
\end{aligned}$$

Uppgift 6

P.g.a. exponentialfördelningens minneslöshet gäller att [se härledningen i Blom et al sid 61]
$P(X > 5 | X > 2) = P(X > 5 - 2) = P(X > 3) = e^{-\lambda \cdot 3} = e^{-3/4} = 0.472$

Uppgift 7

P.g.a. C.G.S. är både ΣX_i och ΣY_i approximativt Normalfördelade. P.g.a. att alla linjärkombinationer av ober Normalfördelade stokastisska variabler är Normalfördelade så är $\bar{Y} - \bar{X}$ det och vi kan använda §12.3 i F.S. och får:

$$I_{\mu_x - \mu_y} = \bar{x} - \bar{y} \pm \sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}} \cdot \lambda_{\alpha/2}$$

Med insatta numeriska värden och $\lambda_{\alpha/2} = \lambda_{0.005} = 2.5758$ fås:

$$I_{\mu_x - \mu_y} = -71 \pm 2.20$$

D.v.s. undre gränsen är -73.20

Uppgift 8

Alternativ A är rätt. Se läroboken.

Uppgift 9

Eftersom vi har två oberoende observationer kommer Likelihoodfunktionen att se ut på följande sätt.

$$L(p) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) = \binom{n_1}{x_1} p^{x_1}(1-p)^{n_1 - x_1} \cdot \binom{n_2}{x_2} p^{x_2}(1-p)^{n_2 - x_2} =$$

$$lnL(p) = ln\left(\binom{n_1}{x_1} \cdot \binom{n_2}{x_2}\right) + (x_1 + x_2)lnp + (n_1 + n_2 - x_1 - x_2)ln(1-p)$$

$$\frac{d}{dp} lnL(p) = \frac{x_1 + x_2}{p} + \frac{n_1 + n_2 - x_1 - x_2}{1-p} \cdot (-1) = 0$$

$$\Rightarrow (1-p)(x_1 + x_2) + p(x_1 + x_2 - n_1 - n_2) = 0$$

$$\Rightarrow x_1 + x_2 - p(n_1 + n_2) = 0$$

$$\Rightarrow p^*_{obs_{ML}} = \frac{x_1 + x_2}{n_1 + n_2}$$

Uppgift 10

$$V(\Theta^*) = V(\frac{1}{kN} \cdot \Sigma X_i) = \frac{1}{k^2 \cdot N^2} V(\Sigma X_i) = [\text{ober}] = \frac{1}{k^2 \cdot N^2} \cdot N \cdot V(X_i) = \frac{1}{k^2 \cdot N} \cdot \frac{k}{\Theta^2} = \frac{1}{kN \cdot \Theta^2}$$

Uppgift 11

Rätt svar är D. Se övnuppg 11.1 i läroboken.

Uppgift 12

P-värdet är P(förkasta H_0) om H_0 är sann $= P(X > Q)$ där $Q \in \chi^2(3-1)$
Tab 4 $\Rightarrow P(X > \chi^2(2) > 7.38) = 0.025$, $P(X > \chi^2(2) > 9.21) = 0.01$
Alltså måste P(X¿7.41) ligga mellan 0.01 och 0.025.
Det enda alternativet som gör det är B $\Rightarrow$ P-värdet = 2.46%.

Del II

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

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

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

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

Enligt lagen om total sannolikhet fås

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

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

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

Uppgift 14

Inför beteckningen X_i för antal bilar tillhörande lägenhet nr i och Y för totala antal bilar, vi har också att

$$\mathrm{Y} = \sum_{i=1}^{80} \mathrm{X}_i = \mathrm{X}_1 + \mathrm{X}_2 + \ldots + \mathrm{X}_{80}.$$

Den sökta sannolikheten är $P(\mathrm{Y} \leq 80)$.
Vi antar att X_i:na är oberoende och likafördelade. Y blir då en summa av många, oberoende och likafördelade variabler vilket leder till att CGS kan användas. $\mathrm{Y} \sim \mathrm{N}(n \cdot \mu, \sqrt{n} \cdot \sigma)$, där $n = 80$, $\mu = E(\mathrm{X}_i)$ och $\sigma = D(\mathrm{X}_i)$

$$\begin{aligned}
E(\mathrm{X}_i) &= \sum_{k=0}^{2} k P(\mathrm{X}_i = k) = 0 \cdot 0.4 + 1 \cdot 0.4 + 2 \cdot 0.2 = 0.8 \\
E\left(\mathrm{X}_i^2\right) &= \sum_{k=0}^{2} k^2 P(\mathrm{X}_i = k) = 0^2 \cdot 0.4 + 1^2 \cdot 0.4 + 2^2 \cdot 0.2 = 1.2 \\
V(\mathrm{X}_i) &= E\left(\mathrm{X}_i^2\right) - E(\mathrm{X}_i)^2 = 1.2 - 0.8^2 = 0.56 \\
D(\mathrm{X}_i) &= \sqrt{V(\mathrm{X}_i)} = \sqrt{0.56}
\end{aligned}$$

Vi får då att Y $\sim$ N$(80 \cdot 0.8, \sqrt{80 \cdot 0.56})$ och den sökta sannolikheten blir

$$P\left(\mathrm{Y} \leq 80\right) = P\left(\frac{\mathrm{Y} - 80 \cdot 0.8}{\sqrt{80 \cdot 0.56}} \leq \frac{80 - 80 \cdot 0.8}{\sqrt{80 \cdot 0.56}}\right) = \Phi(\frac{16}{\sqrt{80 \cdot 0.56}}) \approx$$
$$\approx \Phi(2.39) = 0.99$$

Svar: <u>Antalet bilplatser räcker med sannolikheten 0.99.</u>

Uppgift 15

Två oberoende stickprov med N(μ_A, σ)- respektive N(μ_B, σ)-fördelade observationer. Ett 90%-igt konfidensintervall för $\mu_A - \mu_B$ blir med t-metoden (FS 11.2)

$$\bar{x} - \bar{y} \pm t_{0.05}(7 + 7 - 2)s\sqrt{\frac{2}{7}}$$

där $\bar{x} = 33.72857143$ och $\bar{y} = 21.71428571$. Vidare får vi

$$s_A = 1.165781977$$

$$s_B = 0.7988086367$$

$$s^2 = \frac{(7-1)s_A^2 + (7-1)s_B^2}{7+7-2}$$

som ger

$$s = 0.9992854587$$

och vi får intervallet till $12.104 \pm 1.78 \cdot 0.9992854587\sqrt{2/7} = \underline{12.104 \pm 0.951}$.

Vi tar $H_0 : \mu_A = \mu_B$ och $H_1 : \mu_A \neq \mu_B$. Vi förkastar H_0 på signifikansnivån 10% eftersom konfidensintervallet inte innehåller 0. Slutsatsen är alltså att det spelar roll om man härdar cementen 2 eller 7 dagar.

Uppgift 16

Låt $x_1, x_2, \ldots, x_n$ beteckna data från mätningarna av cirkelns diameter. Dessa är observationer av $X_1, X_2, \ldots, X_n$ som är $N(d, \sigma)$, medan data $y_1, y_2, \ldots, y_m$ från mätningar av cirkelns omkrets är observationer av $Y_1, Y_2, \ldots, Y_m$ som är $N(\pi d, \sigma)$. Samtliga X_i:n och Y_i:n är oberoende av varandra.

1) För att beräkna MK-skattningen av d baserat enbart på observationerna $x_1, x_2, \ldots, x_n$ betraktar vi

$$Q(d) = \sum_{i=1}^{n}(x_i - d)^2$$

Derivering map d ger

$$Q'(d) = \sum_{i=1}^{n} -2(x_i - d)$$

Sätter vi derivatan till 0 fås ekvationen

$$\sum_{i=1}^{n} x_i - nd = 0.$$

Löser man denna för d får man att MK-skattningen av d baserat på observationerna $x_1, x_2, \ldots, x_n$ är

$$\underline{d^*_{obs} = \frac{1}{n} \sum_{i=1}^{n} x_i.}$$

2) För att beräkna MK-skattningen av d baserat enbart på observationerna $y_1, y_2, \ldots, y_m$ betraktar vi

$$Q(d) = \sum_{i=1}^{m} (y_i - \pi d)^2$$

Derivering map d ger

$$Q'(d) = \sum_{i=1}^{m} -2\pi(y_i - \pi d)$$

Sätter vi derivatan till 0 fås ekvationen

$$\sum_{i=1}^{m} y_i - m\pi d = 0.$$

Löser man denna för d får man att MK-skattningen av d baserat på observationerna $y_1, y_2, \ldots, y_m$ är

$$\underline{\widehat{d}_{obs} = \frac{1}{m\pi} \sum_{i=1}^{m} y_i.}$$

3) Slutligen för att beräkna MK-skattningen av d baserat på samtliga observationer, dvs baserat på $x_1, x_2, \ldots, x_n$ och $y_1, y_2, \ldots, y_m$ betraktar vi

$$Q(d) = \sum_{i=1}^{n} (x_i - d)^2 + \sum_{i=1}^{m} (y_i - \pi d)^2.$$

Derivering map d ger

$$Q'(d) = \sum_{i=1}^{n} -2(x_i - d) + \sum_{i=1}^{m} -2\pi(y_i - \pi d)$$

Sätter vi derivatan till 0 fås ekvationen

$$\sum_{i=1}^{n} x_i - nd + \pi \sum_{i=1}^{m} y_i - m\pi^2 d = 0.$$

Löser man denna för d får man att MK-skattningen av d baserat på observationerna $x_1, x_2, \ldots, x_n$ och $y_1, y_2, \ldots, y_m$ är

$$d^{MK}_{obs} = \underline{\frac{\sum_{i=1}^{n} x_i + \pi \sum_{i=1}^{m} y_i}{n + \pi^2 m}}.$$

Vi beräknar nu variansen för den till d^*_{obs} hörande stickprovsvariabeln d^*. Vi får

$$\begin{aligned} V(d^*) &= V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) \\ &= \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}. \end{aligned}$$

Här har vi utnyttjat att X_i:na är oberoende för att få andra likheten och att $X_i \in N(d, \sigma)$ för att få tredje.
Variansen för stickprovsvariabeln svarande mot $\widehat{d}_{obs}$ ges av

$$\begin{aligned} V(\widehat{d}) &= V\left(\frac{1}{\pi m}\sum_{i=1}^{m} Y_i\right) \\ &= \frac{1}{\pi^2 m^2}\sum_{i=1}^{n} V(Y_i) \\ &= \frac{1}{\pi^2 m^2} m\sigma^2 = \frac{\sigma^2}{\pi^2 m}. \end{aligned}$$

Här har vi utnyttjat att Y_i:na är oberoende för att få andra likheten och att $Y_i \in N(\pi d, \sigma)$ för att få tredje.
Om vi nu låter $\sigma_1 = V(d^*)$ och $\sigma_2 = V(\widehat{d})$ så får vi att

$$a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} = \frac{\sigma^2/(\pi^2 m)}{\sigma^2/n + \sigma^2/(\pi^2 m)} = \frac{n}{n + \pi^2 m}.$$

Insatt i formel (1) får vi nu

$$\begin{aligned} \widetilde{d}_{obs} &= \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} d^*_{obs} + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\widehat{d}_{obs} \\ &= \frac{n}{n + \pi^2 m} \cdot \frac{1}{n}\sum_{i=1}^{n} x_i + \frac{\pi^2 m}{n + \pi^2 m} \cdot \frac{1}{\pi m}\sum_{i=1}^{m} y_i \\ &= \frac{1}{n + \pi^2 m}\sum_{i=1}^{n} x_i + \frac{\pi}{n + \pi^2 m}\sum_{i=1}^{m} y_i \end{aligned}$$

och därmed ser vi att $\underline{d^{MK}_{obs} = \widetilde{d}_{obs}}$, dvs MK-skattningen av d baserat på alla $n + m$ mätningar är den effektivaste skattningen man kan få då man kombinerar ihop skattningen baserat på de n mätningarna av cirkelns diameter och de m mätningarna av cirkelns omkrets linjärt.