

Tentamen i matematisk statistik för MI/EPI/DI/MEI den 19 dec 2012

Uppgift 1: Ett företag tillverkar säkerhetsutrustningar till bilar. Tillverkningen är förlagd till fyra olika länder, A, B C och D. I land A finns 40% av företagets produktion, i land B 25%, i land C 20% och i land D den resterande delen. Man vet sedan tidigare att 90% av utrustningarna som kommer från A har en livslängd som är minst 5 år. Motsvarande andelar för B, C och D är 95%, 97% och 99%. Hela företagets tillverkning samlas i ett stort centrallager innan några utrustningar säljs vidare till kund. När en bilfabrik köper företagets säkerhetsutrustningar så får den enheter från alla fyra länderna. Anta att man väljer ut en utrustning slumpmässigt från bilfabrikens lager.

- Vad är sannolikheten att den utvalda utrustningen håller minst 5 år?
- Om säkerhetsutrustningen håller minst 5 år, hur stor är sannolikheten att den kommer från land D?

(6 poäng)

Uppgift 2: Till en grossist levereras bl. a. en kartong med 50 jultomar i porslin. Fyra av jultomtarna har blivit kantstötta i leveransen och måste returneras till fabriken. Ytterligare sex har bara blivit lite skavda så att den röda färgen har försvunnit på någon fläck. Dessa tomtar kan säljas till reducerat pris. Resterande fyrtio tomtar är oskadade. Anta att man slumpmässigt väljer ut fyra jultomtar från leveransen.

- Vad är sannolikheten att man inte får någon tomte, som är kantstött?
- Vad är sannolikheten att man får exakt en tomte som är kantstött och exakt en som är lite skavd?
- Vad är sannolikheten att man får högst en tomte som är kantstött och högst en som är skavd?

(8 poäng)

Uppgift 3: Anta att en fluga sätter sig på en mätsticka, som är graderad mellan 0 och 25 cm.

- Vad är sannolikheten att den sitter närmare markeringen 0 än 25?
- Vad är sannolikheten att flugan sitter minst tre gånger så långt från 0 som från 25?

(6 poäng)

Uppgift 4: När man kastar varpa siktar man på en pinne nedsatt i marken. Varpan (en speciell sorts sten) skall komma så nära pinnen som möjligt. Anta att man kastar varpan utefter en 7 meter lång linje med avsikten att komma så nära pinnen som möjligt. Pinnen identifieras med värdet 0 medan ξ är den faktiska landningspunkten för varpan. Frekvensfunktionen för landningspunkten kan skrivas enligt följande:

$$f(x) = \begin{cases} 0.75(1-x^2) & \text{för } -a \leq x \leq a \\ 0 & \text{för övrigt} \end{cases}$$

Kasta varpan 10 gånger. Beräkna variansen för antal gånger som varpan hamnar närmare pinnen än 0.5m.

(6 poäng)

Uppgift 5: Ett mejeri vill kontrollera densiteten på förpackad vispgrädde och väger därför 196 gräddförpackningar. Vikterna, ξ , (i gram) anses vara oberoende stokastiska variabler med väntevärdet $E(\xi)$ och standardavvikelsen 1.60 gram. Beräkna

$$P(|\bar{\xi} - E(\xi)| < 0.20)$$

där $\bar{\xi}$ är medelvärdet av de 196 gräddförpackningarnas vikter.

(6 poäng)

Uppgift 6: Anta att man genomförde ett faktorförsök där de 3 faktorerna A, B och C användes. I varje försökssituation användes 5 replikat. När ett försök hade pågått tillräckligt länge avbröts det och man användes sig då av de värden som framkommit hittills. Detta hände i försök 3, 5 och 6, vilket gör att dessa värden ser lite konstiga ut. Vid analysen användes ändå dessa värden på samma sätt som de övriga. Följande försöksresultat erhöles:

A	B	C	Medel- värde	Standard- avvikelse
-	-	-	1875	1541.2
+	-	-	1635	1333.4
-	+	-	150	0.0
+	+	-	165	32.6
-	-	+	3000	0.0
+	-	+	3000	0.0
-	+	+	1942	1346.8
+	+	+	2512	1090.3

Följande effekter beräknades

$$\ell_A = 86.25 \quad \ell_B = -1185.25 \quad \ell_{AB} = 206.25 \quad \ell_{ABC} = 78.75$$

- Använd ovanstående för att bestämma de resterande effekterna.
- Beräkna ett 99%-igt referensintervall för effekterna.
- Ange en matematisk modell för ovanstående situation där enbart signifikanta effekter ingår.

(6 poäng)

Uppgift 7: Anta att man vill bestämma en lämplig provtagningsplan som går genom punkterna ($p_1 = 0.01$, $L(p) = 0.95$) och ($p_2 = 0.05$, $L(p) = 0.10$) på OC-kurvan.

- Bestäm en enkel provtagningsplan som uppfyller detta genom att använda det bifogade binomialfördelningsnomogrammet.
- Bestäm en dubbel provtagningsplan som uppfyller detta.

(6 poäng)

Uppgift 8: För en process som är under statistisk kontroll är $p = 3\%$. Man tar dagligen 400 enheter för kontroll och använder p-diagram med 3-sigma-gränser.

- a) Vad är genomsnittligt antal inprickade punkter i kontrolldiagrammet fram till ett falsklarm, ARL?
- b) Vad är sannolikheten att en plötslig ändring till $p = 6\%$ upptäcks
 - 1) vid första provgruppen efter ändringen?
 - 2) vid någon av de tre första provgrupperna efter ändringen?

(6 poäng)

Lösningar till Matematisk statistik den 19/12-12

Uppgift 1: Låt A, B, C respektive D beteckna att säkerhetsutrustningen är tillverkad i land A, B, C respektive D. Låt vidare L betyda att utrustningen har en livslängd på minst 5 år.

$$P(A) = 0.49 \quad P(B) = 0.25 \quad P(C) = 0.20 \quad P(D) = 0.15.$$

$$P(L|A) = 0.90 \quad P(L|B) = 0.95 \quad P(L|C) = 0.97 \quad P(L|D) = 0.99$$

$$\begin{aligned} \text{a) } P(L) &= P(L|A) \cdot P(A) + P(L|B) \cdot P(B) + P(L|C) \cdot P(C) + P(L|D) \cdot P(D) = \\ &= 0.90 \cdot 0.49 + 0.95 \cdot 0.25 + 0.97 \cdot 0.20 + 0.99 \cdot 0.15 = 0.94 \end{aligned}$$

$$\text{b) } P(D|L) = (\text{Bayes sats}) = \frac{P(L|D) \cdot P(D)}{P(L)} = \frac{0.99 \cdot 0.15}{0.94} \approx 0.158$$

Uppgift 2: $N = 50$ $n = 4$ Kantstötta = 4 Skavda = 6 Hela = 40

$$\text{a) } P(\text{ingen kantstött}) = \frac{\binom{4}{0} \binom{46}{4}}{\binom{50}{4}} = \frac{163185}{230300} \approx 0.709$$

$$\text{b) } P(\text{en kantstött och en skavd}) = \frac{\binom{4}{1} \binom{6}{1} \binom{40}{2}}{\binom{50}{4}} = \frac{18720}{230300} \approx 0.08$$

c) Det finns fyra fall som uppfyller villkoret att högst en är kantstött och högst en är skavd. Villkoret "högst en" innebär "noll eller en".

Låt K, S och H beteckna kantstött, skavd respektive hel jultomte. De fyra fallen kan de skrivas som

(K=0, S=0, H=4) och (K=0, S=1, H=3) och (K=1, S=0, H=3) och (K=1, S=1, H=2)

Motsvarande sannolikheter blir

$$P(K=0, S=0, H=4) + P(K=0, S=1, H=3) + P(K=1, S=0, H=3) + P(K=1, S=1, H=2) =$$

$$= \frac{\binom{4}{0} \binom{6}{0} \binom{40}{4}}{\binom{50}{4}} + \frac{\binom{4}{0} \binom{6}{1} \binom{40}{3}}{\binom{50}{4}} + \frac{\binom{4}{1} \binom{6}{0} \binom{40}{3}}{\binom{50}{4}} + \frac{\binom{4}{1} \binom{6}{1} \binom{40}{2}}{\binom{50}{4}} = \frac{208910}{230300} \approx 0.907$$

Uppgift 3: ξ = antal cm från 0:an på en 50 cm lång linjal. ξ är $R(a, b) = R(0, 50)$

a) $P(\text{närmare 0 än 50}) = P(\text{nedre halvan}) = P(\xi < 25) = \frac{25-0}{50-0} = 0.5$

b) $P(\text{minst 3 gånger så långt från 0 som från 50}) = P(\text{övre fjärdedelen av linjalen}) =$
 $= P(\xi > \frac{3 \cdot 50}{4}) = 1 - P(\xi < 37.5) = 1 - \frac{37.5-0}{50-0} = 0.25$

Uppgift 4: ξ = landningspunkten för varpan räknat från pinnen

Bestäm variansen för antal av 10 kast som kommer närmare pinnen än 0.5m.

η = antal kast som kommer närmare pinnen än 0.5m

η är $\text{Bin}(n, p) = \text{Bin}(10, p)$ där p beräknas med hjälp av frekvensfunktionen $f(x)$.

Beräkna värdet på a .

$$0.75 \int_{-a}^a (1-x^2) dx = 0.75 \left[x - \frac{x^3}{3} \right]_{-a}^a = 0.75 \left[\left(a - \frac{a^3}{3} \right) - \left(-a - \frac{(-a)^3}{3} \right) \right] = 1 \Rightarrow 3a - a^3 = 2$$

$\Rightarrow a = 1$ Om a hade varit mindre än 0.5 så hade p blivit 1

$$p = 0.75 \int_{-0.5}^{0.5} (1-x^2) dx = 0.75 \left[x - \frac{x^3}{3} \right]_{-0.5}^{0.5} = 0.75 \left[\left(0.5 - \frac{0.5^3}{3} \right) - \left(-0.5 - \frac{(-0.5)^3}{3} \right) \right] =$$
$$= 0.75 \cdot \left(1 - \frac{0.25}{3} \right) = 0.6875$$

Alltså, η är $\text{Bin}(10, 0.6875) \Rightarrow \text{Var}(\eta) = np(1-p) = 10 \cdot 0.6875 (1 - 0.6875) \approx 2.148$

Uppgift 5: ξ = vikt $S(\xi) = 1.60$ gram $n = 196$ förpackningar

$$E(\bar{\xi}) = E(\xi) \quad \text{Var}(\bar{\xi}) = \text{Var}\left(\frac{\text{Var}(\xi)}{n}\right) = \frac{1.6^2}{196}$$

$$P(|\bar{\xi} - E(\xi)| < 0.20) = P(-0.20 < \bar{\xi} - E(\xi) < 0.20) = P(\bar{\xi} - E(\xi) < 0.20) - P(\bar{\xi} - E(\xi) < -0.20) =$$
$$= P\left(Z < \frac{0.20}{\frac{1.6}{\sqrt{196}}}\right) - (1 - P\left(Z < \frac{0.20}{\frac{1.6}{\sqrt{196}}}\right)) = 2 P(Z < 1.75) - 1 = 0.9198$$

Uppgift 6:

$$a) \ell_C = \frac{3000 + 3000 + 1942 + 2512}{4} - \frac{1875 + 1635 + 150 + 165}{4} = 1657.25$$

$$\ell_{AC} = \frac{1875 + 150 + 3000 + 2512}{4} - \frac{1635 + 165 + 3000 + 1942}{4} = 198.75$$

$$\ell_{BC} = \frac{1875 + 1635 + 1942 + 2512}{4} - \frac{150 + 165 + 3000 + 3000}{4} = 412.25$$

$$b) s_p^2 = \frac{(5-1)1541.2^2 + (5-1)1332.4^2 + \dots + (5-1)1090.3^2}{8 \cdot 4} = 223571.0716$$

df = 32 Eftersom $n > 30 \Rightarrow$ använd normalfördelningsapproximation

Ett 99%-igt referensintervall

$$0 \pm 2.575 \cdot \frac{2 \cdot \sqrt{223571.0716}}{\sqrt{8 \cdot 5}} \Rightarrow 0 \pm 385.02$$

c) Faktorerna B och C samt samspelet BC ger signifikant effekter

$$\hat{y} = \ell_M + \frac{\ell_B}{2} \cdot x_B + \frac{\ell_C}{2} \cdot x_C + \frac{\ell_{BC}}{2} \cdot x_{BC} \Rightarrow$$

$$\hat{y} = 1784.875 - \frac{1185.25}{2} \cdot x_B + \frac{1657.25}{2} \cdot x_C + \frac{412.25}{2} \cdot x_{BC}$$

$$\text{där } \ell_M = \frac{1875 + 1635 + 150 + 165 + 3000 + 3000 + 1942 + 2512}{8} = 1784.875$$

Uppgift 7:

a) Skärningspunkten mellan linjerna ligger inte exakt på ett entydigt värde för n och c utan blir ungefär $n=120$ och $c=3$.

b) Kvoten $p_2/p_1 = 0.05/0.01 = 5$. Detta medför enligt tabellen längst bak i formelsamlingen att vi skall välja provtagningsplan 5 för alternativ 2, dvs $n_2=2n_1$.

Den dubbla provtagningsplan som passar bäst är $c_1 = 1$, $c_2 = 4$ och $r_1 = r_2 = 5$. Vidare finner vi sambandet

$$p_1 n_1 = 0.01 n_1 = 0.77 \Rightarrow p_1 = 1\% \Rightarrow n_1 = \frac{0.77}{0.01} = 77 \Rightarrow n_2 = 2n_1 = 154$$

En lämplig provtagningsplan blir då

$$n_1 = 77, n_2 = 154, c_1 = 1, c_2 = 4 \text{ och } r_1 = r_2 = 5.$$

Uppgift 8: $p = 0.03$ och $n = 400$

a) Om vi använder ett p-diagram *andelen* defekta enheter får vi följande styrgränser

$$S_{\bar{o}} = p + 3 \sqrt{\frac{p(1-p)}{n}} = 0.03 + 0.0256 = 0.0556$$

$$CL = p = 0.03$$

$$S_u = p - 3 \sqrt{\frac{p(1-p)}{n}} = 0.03 - 0.0256 = 0.0044$$

$$\begin{aligned} P(\text{falsklarm}) &= 1 - L(p) = P(\hat{p} < 0.0044) + P(\hat{p} > 0.0556) = P\left(Z < \frac{0.0044 - 0.03}{\sqrt{\frac{0.03(1-0.03)}{400}}}\right) + \\ &+ 1 - P\left(Z < \frac{0.0556 - 0.03}{\sqrt{\frac{0.03(1-0.03)}{400}}}\right) = P(Z < -3) + 1 - P(Z < 3) = 1 - P(Z < 3) + 1 - P(Z < 3) = \\ &= 2 - 2 \cdot 0.9987 = 0.0026 \end{aligned}$$

$$ARL = \frac{1}{1-L(p)} = \frac{1}{0.0026} = 384.61 \approx 385$$

$$\begin{aligned} \text{b)} \quad P(\text{larm}) &= P(\hat{p} < 0.0044) + P(\hat{p} > 0.0556) = P\left(Z < \frac{0.0044 - 0.06}{\sqrt{\frac{0.06(1-0.06)}{400}}}\right) + \\ &+ 1 - P\left(Z < \frac{0.0556 - 0.06}{\sqrt{\frac{0.06(1-0.06)}{400}}}\right) = P(Z < -4.6) + 1 - P(Z < -0.37) = \\ &= 0 + P(Z < 0.37) = 0.6447 \end{aligned}$$

1) $P(\text{larm vid första provgruppen efter ändringen}) = 0.6447$

2) $P(\text{vid någon av de tre första provgrupperna efter ändringen}) =$
 $P(\text{larm vid första provgruppen}) + P(\text{larm vid andra provgruppen}) +$
 $+ P(\text{larm vid tredje provgruppen}) = 0.6447 + 0.3553 \cdot 0.6447 +$
 $+ 0.3553 \cdot 0.3553 \cdot 0.6447 = 0.9466$