Lösningsförslag TMSB18 Matematisk statistik IL 10101
Tid: 12.00-17.00
Telefon: 101620,
Examinator: F Abrahamsson

TEKNISKA HÖGSKOLAN
HÖGSKOLAN I JÖNKÖPING

1. Varje dag levereras en last med 100 maskindetaljer till ett företag. Man tar då ett stickprov och kontrollerar 10 av detaljerna och skickar tillbaka sändningen om man finner mer än 1 felaktig detalj i stickprovet. Om leveransen innehåller 12 felaktiga detaljer, vad är då sannolikheten att den skickas tillbaka? (2p)

   Lösning: Om ξ är antalet felaktiga detaljer i stickprovet så är $\xi \in Hyp(100, 10, 0.12)$. Sannolikheten att sändningen skickas tillbaka är

$$\begin{aligned} P(\xi > 1) &= 1 - P(\leq 1) = 1 - (P(\xi = 0) + P(\xi = 1)) \\ &= 1 - \frac{\binom{12}{0}\binom{88}{10}}{\binom{100}{10}} - \frac{\binom{12}{1}\binom{88}{9}}{\binom{100}{10}} \\ &\approx 0.34 \end{aligned}$$

2. Ett företags nettoinkomster ξ uppfyller

$$\xi = 2\xi_1 - 3\xi_2$$

   där ξ_1, ξ_2 är oberoende stokastiska variabler med $E(\xi_1) = 45$, $E(\xi_2) = 25$ (enhet miljoner kronor) och $V(\xi_1) = 100$, $V(\xi_2) = 75$.

   (a) Beräkna väntevärde och standardavvikelse för företagets nettoinkomster. (2p)

   (b) Om man vet att ξ_1 och ξ_2 är normalfördelade, beräkna då sannolikheten att nettoinkomster är större än 10 miljoner kronor. (1p)

   Lösning: (a)

$$\mu = E(\xi) = 2E(\xi_1) - 3E(\xi_2) = 2 \cdot 45 - 3 \cdot 25 = 15$$

   och

$$V(\xi) = 2^2 V(\xi_1) + (-3)^2 V(\xi_2) = 4 \cdot 100 + 9 \cdot 75 = 1075$$

   så att standardavvikelsen blir

$$\sigma = \sqrt{V(\xi)} = \sqrt{1075} \approx 32.8\,.$$

   (b) Om ξ_1, ξ_2 är normalfördelade så blir även ξ normalfördelad, dvs. $\xi \in N(15, 32.8)$. Då är

$$P(\xi > 10) = 1 - P(\xi \leq 10) = 1 - \Phi\left(\frac{10 - 15}{32.8}\right) = 1 - \Phi(-0.15) = \Phi(0.15) \approx 0.56\,.$$

3. En tillverkare av vitamintabletter påstår att det finns i genomsnitt 47 μg B-vitamin i varje tablett. En konsumentorganisation misstänker att det i själva verket är mindre än 47 μg och analyserar därför vitamininnehållet i 20 tabletter. Från denna analys fås medelvärdet $\overline{x} = 46.3$ μg. Vi antar att B-vitaminmängden i en tablett är normalfördelad med genomsnittligt innehåll μ och känd standardavvikelse $\sigma = 3.5$ μg.

   (a) Formulera och genomför ett lämpligt hypotestest som konsumentorganisationen skulle kunna använda för att bekräfta sina misstankar. Använd signifikansnivån 5%. (2p)

   (b) Vilket P-värde har testet? (1p)

   (c) Om tabletterna i genomsnitt innehåller 45 μg, vilken är då testets styrka? (1p)

   Lösning: (a) Ett lämpligt test är

$$H_0 : \mu = 47$$
$$H_1 : \mu < 47$$

   med signifikansnivån $\alpha = 0.05$. Testvariabeln är

$$Z_0 = \frac{\overline{\xi} - 47}{3.5/\sqrt{20}}$$

och vi förkastar H_0 på 5% signifikansnivå om $z_0 < -\lambda_{0.05} = -1.64$. Med $\overline{x} = 46.3$ fås $z_0 = -0.89$ och vi kan då inte förkasta H_0.

(b) $P-$ värdet $= P(Z_0 \leq -0.89) = \Phi(-0.89) = 1 - \Phi(0.89) \approx 0.19$.

(c) Om $\mu = 45$ så är

$$Z_0 = \frac{\overline{\xi} - 47}{3.5/\sqrt{20}} = \underbrace{\frac{\overline{\xi} - 45}{\sigma/\sqrt{20}}}_{\in N(0,1)} - \frac{2}{3.5/\sqrt{20}} \in N(-\frac{2\sqrt{20}}{3.5}, 1) \approx N(-2.56, 1)\,.$$

Styrkan är därför

$$\begin{aligned} 1 - \beta &= P(Z_0 < -1.64 | Z_0 \in N(-2.56, 1)) = \Phi\left(-1.64 - (-2.56)\right) = \Phi(0.92) \\ &\approx 0.82\,. \end{aligned}$$

4. Kunder anländer till en glasskiosk (med en kassa) enligt en Poissonprocess med intensiteten 15 kunder per timme. Att serva en kund tar i genomsnitt 1.5 minuter.

   (a) Bestäm det förväntade antalet kunder som står i kön till kiosken. Beräkna också den förväntade tiden det tar för en kund att få sin glass. (2p)

   (b) Bestäm sannolikheten att det är högst 2 kunder som står i kön. (1p)

Lösning: (a) I tidsenheten minuter så är ankomstintensiteten $\lambda = 15/60 = \frac{1}{4}$ kunder per minut. Med $1/\mu = \frac{3}{2}$ så är $\mu = \frac{2}{3}$ så att utnyttjandegraden är $\rho = \lambda/\mu = \frac{3}{8}$. Det följer att

$$L = \frac{\rho}{1-\rho} = \frac{3}{5}\,,\ L_q = \frac{\rho^2}{1-\rho} = \rho L = \frac{9}{40}$$

dvs. det förväntade antalet köande kunder är $\frac{9}{40} \approx 0.23$. Littles sats ger sedan att $W = L/\lambda = \frac{12}{5} = 2.4$ minuter.

(b)

$$P(\text{högst 2 kunder}) = \sum_{k=0}^{2} P_k = \sum_{k=0}^{2} \rho^k (1-\rho) = \sum_{k=0}^{2} \left(\frac{3}{8}\right)^k \cdot \frac{5}{8} = \frac{485}{512} \approx 0.95$$

5. En höjdhoppstränare testar en ny träningsmetod som involverar en halvtimme hopprepsträning varje dag. Efter en månad med den nya metoden låter tränaren sina adepter hoppa höjd och jämför resultaten med tidigare resultat (resultaten mäts i cm).

Hoppare	1	2	3	4	5	6
Innan hopprepsträning	168	189	177	199	174	179
Efter hopprepsträning	172	186	183	201	176	181

   (a) Om vi antar att resultaten är normalfördelade, bestäm då ett ensidigt nedåt begränsat 95% konfidensintervall för differensen mellan hoppresultaten efter resp. före hopprepsträningen. Baserat på detta resultat, kan vi hävda att hopprepsträningen har förbättrat resultaten? (2p)

   (b) Om vi inte kan anta att resultatet är normalfördelat, formulera och genomför ett teckentest på 5% nivå för att avgöra om hopprepsträningen har förbättrat resultaten. (1p)

Lösning: (a) Det är här ett parvis ordnat stickprov $(x_1, y_1), \ldots, (x_6, y_6)$ där (x_i, y_i) är hoppresultaten före resp. efter hopprepsträningen. Dessa par är observationer av $\xi_i \in N(\mu_i, \sigma_1)$ resp. $\eta_i \in N(\mu_i + \Delta, \sigma_2)$ med differenser $Z_i = \eta_i - \xi_i \in N(\Delta, \sigma)$ där σ ej är känt. De observerade differenserna beräknas enl. tabellen;

Hoppare	1	2	3	4	5	6
Innan hopprepsträning (x_i)	168	189	177	199	174	179
Efter hopprepsträning (y_i)	172	186	183	201	176	181
Differens $z_i = y_i - x_i$	4	-3	6	2	2	2

Vi beräknar $\overline{z} = 2.17$ och $s = 2.9944$. Enligt teorin är

$$\Delta \geq \overline{z} - t_{0.05}(5) \cdot \frac{s}{\sqrt{6}} = 2.17 - 2.015 \cdot \frac{2.9944}{\sqrt{6}} \approx -0.3$$

och vi kan då inte hävda (med detta resultat som grund) att hopprepsträningen har haft en positiv effekt (vilket skulle kräva att $\Delta \geq k > 0$ för något tal k.)

(b) Hypotestestet är

$$H_0 : \Delta = 0$$
$$H_1 : \Delta > 0$$

och signifikansnivån är $\alpha = 0.05$. Antalet positiva differenser R_+ ger att

$$\begin{aligned} P-\text{värdet} &= P(R_+ \geq 5|H_0) = P(R_+ \geq 5|R_+ \in Bin(6, \frac{1}{2})) \\ &= \left(\begin{array}{c} 6 \\ 5 \end{array}\right)(\frac{1}{2})^5(1-\frac{1}{2})^1 + \left(\begin{array}{c} 6 \\ 6 \end{array}\right)(\frac{1}{2})^6(1-\frac{1}{2})^0 \\ &\approx 0.11 \end{aligned}$$

dvs. vi kan inte förkasta H_0 på 5% signifikansnivå eftersom $P-$värdet är större än α.

6. Antalet förkylningar under ett år visar sig i en viss population vara en stokastisk variabel som är Poissonfördelad med $\lambda = 3$. En ny medicin sänker detta värde till $\lambda = 0.75$ men medicinen fungerar bara för 8 av 10 människor. Man beslutar sig för att ge hela populationen denna medicin. Ett år senare väljs en person på måfå, och det visar sig att denna person har haft en förkylning under året. Vad är sannolikheten att medicinen fungerade för denna person? (3p)

Lösning: Låt A vara händelsen att en individ har haft en förkylning under året och låt B vara händelsen att medicinen fungerade (så att $\lambda = 0.75$). Vi noterar att om $\xi \in Po(\lambda)$ så är $P(\xi = 1) = \lambda e^{-\lambda}$. Bayes sats ger

$$\begin{aligned} P(B|A) &= \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)} \\ &= \frac{0.75e^{-0.75} \cdot 0.8}{0.75e^{-0.75} \cdot 0.8 + 3e^{-3} \cdot 0.2} \\ &\approx 0.905\,. \end{aligned}$$

7. Livslängden för en 10A säkring är exponentialfördelad med en genomsnittlig livslängd om 2.5 år. Om man använder en i taget, bestäm approximativt hur många säkringar man behöver ha i lager av denna typ för att sannolikheten att inte behöva köpa fler säkringar under 50 år skall vara minst 95%. (3p)

Lösning: Låt $\xi = \xi_1 + \cdots + \xi_n$ vara den sammanlagda livslängden för n säkringar, där $\xi_i \in Exp(1/2.5)$ .. Då är $\mu = E(\xi_i) = 2.5$ år och $\sigma = \sqrt{V(\xi)} = 2.5$ år. Om n är hyfsat stort, (vilket vi antar) så ger centrala gränsvärdessatsen att approximativt är $\xi \in N(n\mu, \sqrt{n}\sigma)$ dvs. $\xi \in N(2.5n, 2.5\sqrt{n})$. Sannolikheten att lagret säkringar räcker i 50 år är då

$$P(\xi > 50) = 1 - P(\xi \leq 50) \approx 1 - \Phi\left(\frac{50 - 2.5n}{2.5\sqrt{n}}\right) = 0.95.$$

Omskrivning och användning av identiteten $\Phi(-x) = 1 - \Phi(x)$ ger då att

$$\Phi\left(\frac{2.5n - 50}{2.5\sqrt{n}}\right) = 0.95 \Leftrightarrow \frac{2.5n - 50}{2.5\sqrt{n}} \approx 1.64 \Leftrightarrow 2.5n - 50 = 1.64 \cdot 2.5\sqrt{n}$$

dvs. vi vill lösa rotekvationen $2.5n - 50 = 4.1\sqrt{n}$. Kvadrering av båda sidor ger en andragradsekvation i n med lösningarna 28.8 och 13.9. Prövning i den ursprungliga rotekvationen ger att endast roten $n = 28.8$ är riktig, och alltså är svaret $n = 29$ säkringar.

8. Dimensionering av säkerhetslager för t.ex. beställningspunktsystem har behandlats i kursen. En viktig metod som behandlats brukar ofta betecknas som SERV1 och baseras på att en servicenivå bestäms vilken i sin tur kan användas för att beräkna ett säkerhetslager. Dimensioneringen kan baseras på osäkerhet i efterfrågan, ledtid eller både efterfrågan och ledtid.

(a) Härled uttrycket för dimensionering av säkerhetslager enligt SERV1 givet att det finns en osäkerhet i efterfrågan och att beställningspunktsystem används. Antag att efterfrågan är normalfördelad med väntevärde μ st/vecka och standardavvikelse σ st/vecka. Återfyllnadsledtiden är känd och satt till L veckor. (2p)

(b) Förklara på vilket sätt som egenskaper för oberoende stokastiska variabler användes i deluppgift (a). Ange på vilket sätt som resultatet skulle påverkas (ange intervall för standardavvikelsen över ledtid) om oberoende inte kunde ha antagits. (1p)

(c) Beräkna säkerhetslagret enligt SERV1 för 95% servicenivå (vilken kan antas gälla för alla artiklar i verksamheten) om efterfrågan $D \in N(100, 20)$ och ledtiden L är 3 veckor. Vilken servicenivå skulle detta ge mot kund, i termer av kundorder, om kund typiskt har tre orderrader på varje kundorder. (1p)

Lösning 8

Lösningsförslag

a) SERV1 definieras som sannolikheten att brist uppstår under en ordercykel (en ordercykel är tiden från att en beställning läggs tills dess att leverans är mottagen, vilket tar en ledtid L). Brist undviks om efterfrågan under ledtid är mindre beställningspunkten BP. Om SERV1 får beteckna servicenivån så gäller då att $\mathrm{SERV1} = P\left(\widetilde{D}_L \leq \mathrm{BP}\right)$ där $\widetilde{D}_L$ betecknar den osäkra efterfrågan under ledtiden. Då gäller att: $P\left(\widetilde{D}_L \leq \mathrm{BP}\right) = P\left(\widetilde{D}_L \leq \mu_L + \mathrm{SS}\right)$, där SS står för säkerhetslagret. För att kunna använda den standardiserade normalfördelningstabellen så görs den vanliga omskrivningen: $P\left(\widetilde{D}_L \leq \mu_L + \mathrm{SS}\right) = P\left(\frac{\widetilde{D}_L - \mu_L}{\sigma_L} \leq \frac{\mu_L + \mathrm{SS} - \mu_L}{\sigma_L}\right) = \Phi\left(\frac{\mathrm{SS}}{\sigma_L}\right)$. Det här innebär att $\mathrm{SERV1} = \Phi\left(\frac{\mathrm{SS}}{\sigma_L}\right)$ men det som söks är säkerhetslagret vilket då kan beräknas genom att ta inversfunktionen av fördelningsfunktionen:
$\frac{\mathrm{SS}}{\sigma_L} = \Phi^{-1}(\mathrm{SERV1}) \rightarrow \mathrm{SS} = \Phi^{-1}(\mathrm{SERV1}) \cdot \sigma_L$.
Det var dock inte σ_L som var givet utan σ. Från uppgiften erhålls att återfyllnadsledtiden är L veckor vilket innebär att $\sigma_L = \sqrt{L} \cdot \sigma$. Säkerhetslagret kan då beräknas som:
$\mathrm{SS} = \Phi^{-1}(\mathrm{SERV1}) \cdot \sqrt{L} \cdot \sigma$

b) Antagandet om oberoende stokastiska variabler användes vid beräkning av σ_L som baseras på summan av L st veckoprognoser. Om veckoprognoserna kan antas vara oberoende av varandra så kan varianserna adderas för L veckor vilket ger relationen $\sigma_L = \sqrt{L} \cdot \sigma$.
Om det finns beroenden mellan perioderna skulle σ_L bli större och ligga någonstans i intervallet $\sqrt{L} \cdot \sigma \leq \sigma_L \leq L \cdot \sigma$.

c) För en artikel kan säkerhetslagret enligt deluppgift a beräknas som:
$\mathrm{SS} = \Phi^{-1}(\mathrm{SERV1}) \cdot \sqrt{L} \cdot \sigma = 1{,}64 \cdot \sqrt{3} \cdot 20 = 56{,}81 \approx 57$ st.
För servicenivå på kundorder gäller att brist uppstår om minst en av raderna innebär brist. $\mathrm{SERV1_{Order}} = (\,\mathrm{SERV1_{Orderrad}})^3 = (0{,}95)^3 = 0{,}857$. Alltså innebär en servicenivå på 95% på artikel (orderrad) att servicenivån på order blir ca 86%, vilket också är servicenivån som kund upplever.

Poängfördelning enligt rättningsmall

a) 0,5 poäng för att SERV1 innebär servicenivå i termer av sannolikhet att ingen brist uppstår under en ordercykel, 1 poäng för härledning av SS som funktion av σ_L och 0,5 poäng för beräkning av σ_L.

b) 0,5 poäng för var antagandet görs och 0,5 poäng för intervallet.

c) 0,5 poäng för bestämning av säkerhetslager och 0,5 poäng för servicenivå på order.

Referenser

a) Presentationsunderlag: Ti2 – Säkerhetslagerdimensionering

b) Presentationsunderlag: Ti3 – Aggregering och
Presentationsunderlag: Ti7 – Beroendeanalys

c) Presentationsunderlag: Ti2 – Säkerhetslagerdimensionering