Lösningsförslag LMA522, 18 mars 2019

1. Uppgift 1 (2+3+4 poäng) Vid en trafikkontroll passerar det i genomsnitt fyra bilar per minut. Antal bilar som passerar per minut är Poissonfördelat Po(4), d.v.s. $\lambda = 4$.

   (a) (Vad är sannolikheten att det under en minut kommer in minst två bilar?) Låt ξ vara antal passerande bilar per minut. Då är $\xi \in$ Po(4). Sätt händelsen $\{\xi \geq 2\} = A$. $\lambda = 4$/min. Sökt sannolikhet är

   $$P(\xi \geq 2) = 1 - P(\xi \leq 1) = 1 - e^{-4}\left[\frac{4^0}{0!} + \frac{4^1}{1!}\right] \approx 0.91$$

   (b) (Under en given ett enminutersintervall passerade det minst två bilar. Vad är sannolikheten att det totalt passerade minst tre bilar under samma intervall?). Sätt händelsen $B = \{\xi \geq 3\} = B$. Då söks sannolikheten

   $$P(B|A) = \frac{P(\xi \geq 3)}{P(\xi \geq 2)} = \frac{P(B)}{P(A)} = \{P(B) = P(A) - e^{-4}4^2/2!\} =$$

   $$\frac{1 - e^{-4}\left[\frac{4^0}{0!} + \frac{4^1}{1!} + \frac{4^2}{2!}\right]}{1 - e^{-4}\left[\frac{4^0}{0!} + \frac{4^1}{1!}\right]} = 0.838704 \approx 0.84$$

   (c) Sannolikheten att det passerar fler än 250 bilar under en timme kan beräknas m.h.a. $\zeta := \sum_{k=1}^{60} \xi_k$, där $\xi_k \in$ Po(4.0). Då gäller att $\zeta \in$ Po($4.0 \cdot 60$). Nu kan denna fördelning approximeras med N(μ, σ), där $\mu = 240$ och $\sigma = \sqrt{240}$.

   $$P(\zeta \geq 250) \approx 1 - \Phi\left(\frac{250 - 240}{\sqrt{240}}\right) = 0.26\,.$$

   (Det går också bra att använda centrala gränsvärdessatsen.)

2. Uppgift 2 (2+3 poäng) Resultatet efter 6 mätningar av en normalfördelad stokastisk variabel, gav

   $$25.0, 26.0, 23.5, 23.5, 24.0, 25.0\,.$$

   (a) Ett tvåsidigt 95%:s konfidensintervall för väntevärdet μ om standardavvikelsen σ skattas med $s = 1.0$ ges av

   $$\left(\overline{x} - t_{0.025}(5) \cdot \frac{\sigma}{\sqrt{n}}, \overline{x} + t_{0.025}(5) \cdot \frac{\sigma}{\sqrt{n}}\right) = \left(24.5 - 2.57 \cdot \frac{1}{\sqrt{6}}, 24.5 + 2.57 \cdot \frac{1}{\sqrt{6}}\right) = (23.45, 25.55).$$

   (b) Ett uppåt begränsat 95%:s konfidensintervall för standardavvikelsen σ, om standardavvikelsen skattas till $s = 1.0$ och $\chi^2_{0.95}(5) = 1.15$.

   $$\left[0, \sqrt{\frac{(n-1)s^2}{\chi^2_{0.95}}}\right) = \left[0, \sqrt{\frac{5 \cdot 1.0^2}{1.15}}\right) = [0, \sqrt{4.35}) = [0, 2.09).$$

3. Uppgift 3 (3+2 poäng) Registreringsnummer för fordon ges av tre bokstäver som följs av tre siffror. Antal bokstäver som används är 23.

   (a) Vi beräknar först antalet sätt att bilda en kombination av tre bokstäver där man har exakt två stycken A och ett B. Detta antal ges av $\binom{3}{2} = 3$ (de tre sätten är: AAB, ABA, BAA). Genom att byta ut B mot andra bokstäver som inte är A ser vi att det finns $3 \times 22 = 66$ sätt att bilda en kombination av tre bokstäver där man har exakt två A. Eftersom det finns 23 bokstäver blir antalet kombinationer av tre bokstäver där en bokstav förekommer exakt två gånger lika med $66 \times 23 = 1518$. Till sist blir alltså totala antalet registreringsskyltar $1518 \times 10^3 = 1518000$.

   (b) Sannolikheten p att ett fordon med registreringsskylt har ett nummer som (a) är

   $$\frac{\text{antalet skyltar som i (a)}}{\text{totala antalet skyltar}} = \frac{1518000}{23^3 \times 10^3} = 0.12.$$

4. Uppgift 4 (4p). Låt A vara händelsen att skickat tecken är en 0:a, och låt B vara händelsen att mottaget tecken är en 0:a. Vi vet att $P(A) = 0.45$, $P(B) = 0.4$ samt att $P(B^c|A) = 0.03$. Vi söker $P(A^c|B^c)$. Vi har att

$$P(A^c|B^c) = 1 - P(A|B^c) = 1 - \frac{P(B^c|A)P(A)}{P(B^c)} = 1 - \frac{0.03 \times 0.45}{1 - 0.4} = 0.9775.$$

5. Uppgift 5 (2+2+3 poäng)

   Låt ξ_k =livslängd för komponent k, $k = 1, 2$. Det gemensamma väntevärdet är $\mu := 2.0$ (år) så att $\lambda_1 = \lambda_2 = \lambda = 1/2$.

   (a)

$$P(\text{ båda funkar efter ett år}) = P(\xi_1 \geq 1 \text{ och } \xi_2 \geq 1) = P(\xi_1 \geq 1)P(\xi_2 \geq 1)$$
$$= \left(\int_1^\infty \frac{e^{-t/2}}{2}\right)^2 = e^{-1/4},$$

   där vi använde oberoende i andra likheten.

   (c) Här löser vi uppgift c först och använder den sedan för att lösa uppgift b. Men det går bra att lösa uppgift b först också, utan att använda c.

   Elsystemet fungerar om minst en av komponenterna fungerar. Låt ζ vara systemets livslängd. Om systemet har gått sönder vid tiden t så har båda resistorerna gått sönder vid tiden t. Därför får vi att

$$F(t): \quad = \quad P(\zeta \leq t) = P(\xi_1 \leq t \text{ och } \xi_2 \leq t) = \{\text{ober.}\} = (1 - e^{-\lambda\, t})^2 = (1 - e^{-t/2})^2 \Longrightarrow$$
$$f(t) \quad = \quad F'(t) = 2\lambda(1 - e^{-\lambda t})e^{-\lambda t} = 2\lambda(e^{-\lambda t} - e^{-2\lambda t}) = e^{-t/2} - e^{-t}.$$

   (b) Sannolikheten att systemet fungerar minst ett år är

$$1 - F(1.0) = 1 - (1 - e^{-1/2})^2 = 0.85.$$

6. (1 + 1 + 3 + 3 poäng)

   (a) Se tabell 1

   (b) Se tabell 1

Nr.	A	B	C	AB	AC	BC	ABC
1	+	+	+	+	+	+	+
2	-	+	+	-	-	+	-
3	+	-	+	-	+	-	-
4	-	-	+	+	-	-	+
5	+	+	-	+	-	-	-
6	-	+	-	-	+	-	+
7	+	-	-	-	-	+	+
8	-	-	-	+	+	+	-

Tabell 1: Ifylld tabell för faktorförsöken. Tecknen för sampselsefekterna behöver inte fyllas i.

   (c) Sortera effekterna efter storlek (M är inte intressant då det inte är någon effekt av faktorerna). Räkna ut sannolikhetsvärdet för varje sorterat x-värde.

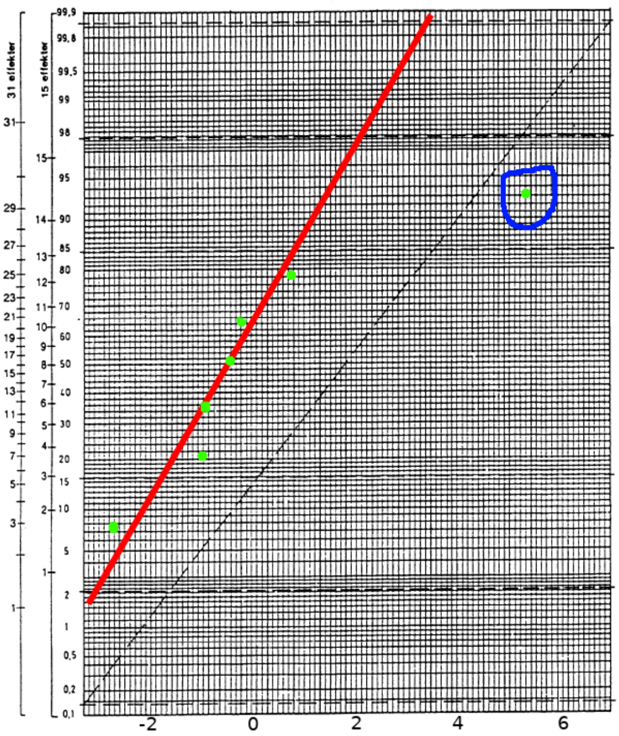
$$p_i = \frac{i - 0.5}{7} \tag{1}$$

   Tabell 2 visar de sorterade effekterna och deras motsvarande sannolikhetsvärde.

   Figur 1 visar det ifyllda normalfördelningsdiagrammet. Som synes, huvudeffekten A är den enda tydligt signifikanta effekten. Det verkar alltså som antal städer besökta är det enda som spelar roll för hur nöjda vännerna är med sina resor.

Effekt	l_C	l_{AC}	l_{BC}	l_{AB}	l_B	l_{ABC}	l_A
x_i	-2.69	-0.97	-0.95	-0.41	-0.20	0.75	5.29
p_i	7.14%	21.4%	35.7%	50%	64.29%	78.57%	92.86%

Tabell 2: Skattade effekter från faktorförsöket.



Figur 1: Normalfördelningsdiagram med de skattade effekterna ifyllda. Huvudeffekten A är omgiven av en rund blå ring för att markera att den var tydligt separerad från det röda strecket.

(d) Eftersom kolumnen för D är identisk med kolumnen för BC så är generatorn, $D = BC$. Detta ger ordet $I = BCD$ och upplösningen blir III.

7. (2 + 3 poäng)

(a) Vi vill skapa ett felantalsdiagram, c-diagram. Eftersom vi vet att om processen skulle vara under statistisk kontroll så kommer antal stulna plånböcker under en fem-veckorsperiod vara fördelad som en Poissonfördelning med parameter $\lambda = 4$. Väntevärde för en sådan fördelning är λ och variansen är också λ. Felantalsdiagrammet får då styrgränserna

$$C_l = \lambda = 4$$
$$S_{ö} = C_l + 3\sqrt{\lambda} = 10$$
$$S_u = \max\{C_l - 3\sqrt{\lambda}, 0\} = 0$$

Ingen av de 10 veckornas mätningar har gett ett antal stulna plånböcker som är orimligt stort eller litet eftersom alla är inom intervallet $[0, 10]$.

(b) Eftersom vi nu är intresserad av antalet stulna plånböcker per vecka och inte per 5 veckor så får vi skala om λ, alltså $\hat{\lambda} = \frac{4}{5} = 0.8$. Antalet stulna plånböcker per

vecka kommer alltså följa en Poisson fördelning med parameter $\hat{\lambda} = 0.8$
Duglihetsindex är

$$C_p = \frac{T_{ö} - T_u}{6\sigma} = \frac{T_{ö} - T_u}{6\sqrt{\hat{\lambda}}} = \frac{3 - 0}{6 \cdot \sqrt{0.8}} \approx 0.56.$$

Då C_p bör vara större än 1.33 så ser vi redan nu att spridningen är för stor för att vi skall vara nöjda med resorna. Eftersom det frågas om korrigerat duglihetsindex så fortsätter vi ändå att räkna ut detta,

$$CM = 2\frac{|\mu - M|}{T_{ö} - T_u} = 2\frac{|\hat{\lambda} - M|}{T_{ö} - T_u} = 2\frac{0.8 - 1.5}{3} = \frac{1.4}{3} \approx 0.47.$$

Det korrigerade duglighetsindexet blir därför

$$C_{pk} = C_p(1 - 0.47) = 0.56 \cdot 0.53 \approx 0.30.$$

Eftersom $0.3 < 1.33$ så lever inte resorna upp till de krav som resebyrån sätter. De måste alltså se över hur de skall kunna öka säkerheten på de resor de säljer.

8. (2 + 3 + 1 + 1 poäng)

(a) Kvoten $\frac{p_2}{p_1} = 5$. Det närmaste värdet är provtagningsplan nr. 5 från tabellen för $2n_1 = n_2$. Detta ger $c_1 = 1, c_2 = 4$ och $r_1 = r_2 = 5$.
Vi ser att $n_1 \cdot 0.02 = 0.77 \Leftrightarrow n_1 = \frac{0.77}{0.02} = 38.5$. Vi ser också att $n_1 \cdot 0.1 = 3.92 \Leftrightarrow n_1 = \frac{3.92}{0.1} = 39.2$. Därför väljer vi $n_1 = 39$ som är heltalet mittemellan 38.5 och 39.2. Vi har alltså den dubbla provtagningsplanen

$$c_1 = 1, c_2 = 4, r_1 = r_2 = 5, n_1 = 39, n_2 = 78.$$

(b) Vi kan approximera antalet funna defekta med binomialfördelningar eftersom $n_1 + n_2 = 117 < 0.1 \cdot 2000$. Antalet defekta i första urvalet kan då modelleras med en slumpvariabel $\xi_1 \sim Bin\left(n = 39, p = \frac{50}{2000} = 2.5\%\right)$ och för andra urvalet med $\xi_2 \sim Bin(n = 78, p = 2.5\%)$.
Acceptanssannolikheten kan räknas ut som

$$\begin{aligned}\mathbb{P}(\text{acceptera}) &= \mathbb{P}(\text{acceptera i urval I}) + \mathbb{P}(\text{acceptera i urval II}) \\ &= \mathbb{P}(\xi_1 \le 1) + \mathbb{P}((\xi_2 \le 4 - \xi_1) \cap (2 \le \xi_1 \le 4)) \\ &= \mathbb{P}(\xi_1 \le 1) + \sum_{i=2}^{4} \mathbb{P}(\xi_2 \le 4 - i)\mathbb{P}(\xi_1 = i).\end{aligned}$$

För att ta fram dessa sannolikheter behöver vi först räkna ut sannolikheten att ξ_1 antar värdena $0, 1, 2, 3, 4$ samt sannolikheterna att ξ_2 antar värdena $0, 1, 2$.

$$\begin{aligned}\mathbb{P}(\xi_1 = 0) &= \binom{39}{0} 0.025^0 \cdot 0.975^{39} = 37.26\% \\ \mathbb{P}(\xi_1 = 1) &= \binom{39}{1} 0.025^1 \cdot 0.975^{38} = 37.26\% \\ \mathbb{P}(\xi_1 = 2) &= \binom{39}{2} 0.025^2 \cdot 0.975^{37} = 18.15\% \\ \mathbb{P}(\xi_1 = 3) &= \binom{39}{3} 0.025^3 \cdot 0.975^{36} = 5.74\% \\ \mathbb{P}(\xi_1 = 4) &= \binom{39}{4} 0.025^4 \cdot 0.975^{35} = 1.33\%\end{aligned}$$

$$\begin{aligned}\mathbb{P}(\xi_2 = 0) &= \binom{78}{0} 0.025^0 \cdot 0.975^{78} = 13.88\% \\ \mathbb{P}(\xi_2 = 1) &= \binom{78}{1} 0.025^1 \cdot 0.975^{77} = 27.76\% \\ \mathbb{P}(\xi_2 = 2) &= \binom{78}{2} 0.025^2 \cdot 0.975^{76} = 27.40\%\end{aligned}$$

Sannolikheten att acceptera i första urvalet är

$$\mathbb{P}(\xi_1 \leq 1) = \mathbb{P}(\xi_1 = 0) + \mathbb{P}(\xi_1 = 1) = 74.52\%.$$

Sannolikheten att acceptera i andra urvalet är

$$\begin{aligned}
&(\mathbb{P}(\xi_2 = 0) + \mathbb{P}(\xi_2 = 1) + \mathbb{P}(\xi_2 = 2))\,\mathbb{P}(\xi_1 = 2)+\\
&(\mathbb{P}(\xi_2 = 0) + \mathbb{P}(\xi_2 = 1))\,\mathbb{P}(\xi_1 = 3) + \mathbb{P}(\xi_2 = 0)\mathbb{P}(\xi_1 = 4)\\
&= 0.6904 \cdot 0.1815 + 0.4164 \cdot 0.0574 + 0.2740 \cdot 0.0133 = 15.29\%
\end{aligned}$$

Slår man ihop dem får man en sannolikhet på 89.81% att acceptera ett parti om det egentligen finns 50 snören som inte tål hårda ryck.

(c) De frågar alltså efter genomsnittsligt provuttag (ASN).

$$\begin{aligned}
ASN(p = 2.5\%) &= n_1 + n_2\mathbb{P}(2 \leq \xi_1 \leq 4)\\
&= 39 + 78 \cdot (0.1815 + 0.0574 + 0.0133) = 58.67.
\end{aligned}$$

(d) Sannolikheten att en slumpvist vald enhet är trasig efter ett parti har gått igenom acceptanskontrollen.