Lösningar till tentamen i Matematisk Statistik, 5p

LGR98

27 oktober, 2001 kl. 9.00 – 13.00

Kursansvarig: Eric Järpe
Maxpoäng: 30
Betygsgränser: 12p: G, 22p: VG
Hjälpmedel: Miniräknare samt tabell- och formelsamling som medföljer tentamenstexten.

Till uppgifterna skall *fullständiga lösningar* lämnas. Lösningarna skall vara *utförligt* redovisade! Varje lösning skall börja överst på nytt papper. Endast en lösning per blad. Lösningar fås med den rättade tentamen då den kvitteras ut.

1. På en skolgård spelas det fotboll på matrasterna. Under en vecka är deltagarantalet

mån	tis	ons	tor	fre
18	8	13	20	19

a) Vad är medelvärdet? ($\frac{1}{2}$p)
b) Vad är standardavvikelsen? ($\frac{1}{2}$p)
c) Rita ett histogram med klasserna $[6, 11]$, $[12, 17]$, $[18, 23]$. (1p)

Lösning:

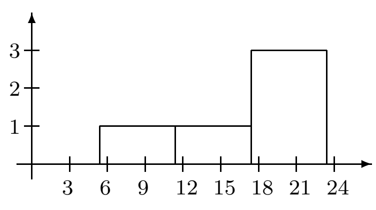
a) $\bar{x} = \frac{1}{5}(18 + 813 + 20 + 19) = 15.6$

b)

$$\begin{aligned} s^2 &= \frac{1}{5-1}\left(\sum_{i=1}^{5} x_i^2 - 5 \cdot (15.6)^2\right) \\ &= \frac{1}{4}\big(1318 - 1216.8\big) \\ &= 25.3 \end{aligned}$$

$$\begin{aligned} s &= \sqrt{25.3} \\ &= 5.03 \end{aligned}$$

c)



2. Antag att man kastar 2 tärningar. Låt A vara händelsen att första tärningen visar 2, B vara händelsen att första tärningen visar 4 eller mindre och C vara händelsen att summan av tärningarna blir 7.
a) Vad är $P(B \cup C)$? (2p)
b) Är A och C oberoende? (2p)
c) Vad är $P(A \mid B)$? (2p)

Lösning:

a) $P(B) = \dfrac{\#\{1,2,3,4\}}{\#\{1,2,3,4,5,6\}} = \dfrac{4}{6} = \dfrac{2}{3}$

$P(C) = \dfrac{\#\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}}{(\#\{1,2,3,4,5,6\}) \cdot (\#\{1,2,3,4,5,6\})} = \dfrac{6}{36} = \dfrac{1}{6}$

$P(B \cap C) = \dfrac{\#\{(1,6),(2,5),(3,4),(4,3)\}}{36} = \dfrac{4}{36} = \dfrac{1}{9}$

$P(B \cup C) = P(B) + P(C) - P(B \cap C) = \frac{2}{3} + \frac{1}{6} - \frac{1}{9} = \frac{6 \cdot 2 + 3 \cdot 1 - 2 \cdot 1}{18} = \frac{8}{18} = \boxed{\frac{4}{9}}$

b) A och C oberoende om $P(A \cap C) = P(A)P(C)$:

$P(A) = \dfrac{\#\{2\}}{\#\{1,2,3,4,5,6\}} = \dfrac{1}{6}$ och $P(C) = \dfrac{1}{6}$ (uträkning av $P(C)$, se ovan)

$P(A \cap C) = \dfrac{\#\{(2,5)\}}{36} = \dfrac{1}{36}$

$P(A \cap C) = \frac{1}{36}$ och $P(A) \cdot P(C) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

så svaret är <u>**ja**, A och C är oberoende.</u>

c) $P(A|B) = P(A \cap B)/P(B)$

$P(A \cap B) = P(A) = \frac{1}{6}$

$P(B) = \frac{2}{3}$ (uträkning, se ovan)

$P(A|B) = \dfrac{1/6}{2/4} = \dfrac{1}{6} \cdot \dfrac{3}{2} = \boxed{\dfrac{1}{4}}$

3. Ylva och Lisa är 2 respektive 7 år gamla. Antag att deras livslängder är oberoende och normalfördelade med väntevärde 75 och standardavvikelse 8. Vad är sannolikheten att Ylva dör senare än Lisa? (3p)

Lösning:
Låt X = Ylvas livslängd $\in N(75, 8)$ och Y = Lisas livslängd $\in N(75, 8)$.

Observera att frågan gäller sannolikheten att Ylva dör senare än Lisa. Eftersom Ylva är 5 år yngre än Lisa är frågan huruvida Ylvas livslängd blir större än Lisas minus 5 år, dvs om $X > Y - 5$.

$P(X > Y - 5) = P(X - Y > -5)$

$E(X - Y) = 75 - 75 = 0$

Eftersom X och Y är oberoende är $V(X - Y) = V(X) + V(Y) = 8^2 + 8^2 = 128$ varmed $X - Y \in N(0, \sqrt{128})$ så

$$\begin{aligned}
P(X > Y - 5) &= P\left(\frac{X - Y - 0}{\sqrt{128}} > \frac{-5 - 0}{\sqrt{128}}\right) \\
&= 1 - P\left(\tfrac{X-Y}{\sqrt{128}} \le \tfrac{-5}{\sqrt{128}}\right) \\
&= 1 - \Phi\left(-\tfrac{5}{\sqrt{128}}\right) \\
&= \Phi\left(\tfrac{5}{\sqrt{128}}\right) \\
&= \Phi(0.442) \\
&= \boxed{0.67}
\end{aligned}$$

4. En gymnastiklärare har en klass på 20 elever. Han observerar att under höstterminen är deltagarantalet på lektionerna

$$20,\ 20,\ 19,\ 19,\ 18,\ 16,\ 19,\ 20$$

Gymnastikläraren frågar matematikläraren (d.v.s. dig) om man kan vara 99% säker på att det i genomsnitt är mindre än 3 frånvarande per gång. Vad svarar du honom? (3p)

Lösning:
Vi vill besvara frågan "är mindre än 3 frånvarande (av 20)?"
Låt X = antal frånvarande. Då har vi observationer av X genom att dra närvarosiffran från 20:
$20-20=0,\ 20-20=0,\ 20-19=1,\ 1,\ 2,\ 4,\ 1,\ 0$
Vi försöker nu visa att det är mindre än 3 frånvarande med osäkerhet 1%,
dvs testa hypotesen

$$\begin{cases} H_0: \mu = 3 \\ H_1: \mu < 3 \end{cases} \quad \text{på sign.nivå } \alpha = 0.01 \text{ där } \mu = E(X).$$

(Alternativt hade vi kunnat låta Y vara antal närvarande, använt närvarosiffrorna såsom de angivits, men då hade vi fått formulera frågan "är antalet närvarande större än 17 (av 20)?" och för besvara denna formulera hypotesen $H_0: \mu = 17$ mot $H_1: \mu > 17$ där $\mu = E(Y)$.)

$$\bar{x} = \frac{9}{8} \qquad s^2 = \frac{1}{7}\left((1^2+1^2+2^2+4^2+1^2) - 8\cdot\left(\frac{8}{9}\right)^2 \right) = 1.84$$

$$U = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{1.125-3}{1.36/\sqrt{8}} = -3.91$$

och $|-3.91| = 3.91 > t_{0.01}(7) = 3$
så nollhypotesen förkastas på signifikansnivå 1%, dvs risken att nollhypotesen felaktigt förkastats är 1%, dvs chansen att nollhypotesen korrekt förkastats är 99%.

Så svaret är **ja**, man kan med 99% säkerhet påstå att färre än 3 elever är frånvarande i genomsnitt.

5. Vid en skola blev alla 405 eleverna på lektionen i svenska erbjudna att välja mellan uppsatsämnena A, B och C. Resultatet blev att 48.9% skrev om A, 23.9% skrev om B och 27.2% skrev om C. Kan man säga att något av ämnena var favorit hos eleverna eller kan det vara bara slumpmässig avvikelse? Gör ett test på lämplig signifikansnivå. (3p)

Lösning:
405 elever varav

	Andel	Obs. antal
A	48.9%	198.045
B	23.9%	96.795
C	27.2%	110.16

Om inget av ämnena var favorit (nollhypotesen) borde det sett ut

	Andel	Obs. antal
A	33.3%	134.865
B	33.3%	134.865
C	33.3%	134.865

Vi vill testa $\begin{cases} H_0 : F = F_0 \\ H_1 : F \neq F_0 \end{cases}$ där F_0 står för fördelningsfunktionen då alla alternativen (A, B och C) är lika sannolika. Testet är chi-två-testet (välj t.ex. $\alpha = 0.05$)

$$\begin{aligned} U &= \sum_{i=1}^{3} \frac{(O_i - E_i)^2}{E_i} \\ &= \frac{(198.045 - 134.865)^2}{134.865} + \frac{(96.795 - 134.865)^2}{134.865} + \frac{(110.16 - 134.865)^2}{134.865} \\ &= 29.6 + 10.6 + 4.5 \\ &= 44.5 \end{aligned}$$

$\chi^2_{0.05}(3-1) = 5.99 < 44.5$ så förkasta H_0 dvs ämnena var **ej** likvärdiga.

6. En familj arrangerar sin trädgårdsbelysning på följande sätt: man seriekopplar en lampa A med en grupp av 5 lampor B_1, B_2, B_3, B_4, B_5. De 5 B-lamporna är parallellkopplade med varandra. Låt livslängden mätt i dagar hos lampa A vara $X_A \in \text{Exp}(0.01)$, hos lampa B_1 vara $X_{B_1} \in \text{Exp}(0.03)$, ..., hos lampa B_5 vara $X_{B_5} \in \text{Exp}(0.03)$ oberoende av varandra. Vad är sannolikheten att lampa A och exakt 2 av de 5 B-lamporna lyser efter 70 dagar? (4p)

Lösning:

För att A-lampan och exakt 2 av B-lamporna ska lysa ska t.ex.
A lysa och av B-lamporna ska nr 1 och 2 lysa och nr 3, 4 och 5 ej lysa.
Låt oss kalla denna händelse $(*)$.

Låt X_A = livslängden hos lampa A
och X_B = livslängden hos lampa B_i där $i = 1, 2, 3, 4, 5$.
Eftersom lamporna lyser oberoende av varandra är

$$\begin{aligned}
P(*) &= P(A \text{ lyser}, B_1, B_2 \text{ lyser}, B_3, B_4, B_5 \text{ släckta}) \\
&= P(X_A > 70)P(X_{B_1} > 70)P(X_{B_2} > 70)P(X_{B_3} \leq 70)P(X_{B_4} \leq 70)P(X_{B_5} \leq 70) \\
&= (1 - (1 - e^{-0.01 \cdot 70}))(1 - (1 - e^{-0.03 \cdot 70}))^2(1 - e^{-0.03 \cdot 70})^3 \\
&= e^{-0.7}e^{2.1 \cdot 2}(1 - e^{-2.1})^3 \\
&= e^{-4.9}(1 - e^{-2.1})^3 \\
&= 0.00503
\end{aligned}$$

Emellertid kan det också hända att
A lyser och av B-lamporna lyser nr 1 och 3 och nr 2, 4 och 5 lyser ej
och denna händelse har samma sannolikhet har samma sannolikhet som $(*)$. Närmare bestämt finns
$\binom{5}{2} = \frac{5 \cdot 4}{2 \cdot 1} = 10$ sådana händelser.
Därmed är slutligen
$P(A \text{ och } 2\ B\text{-lampor lyser}) = 10 \cdot 0.00503 \approx \boxed{0.05}$.

7. Antag att $X \in \text{Po}(1)$ och $Y \in \text{Po}(2)$ och att X och Y är oberoende. Visa att $X + Y \in \text{Po}(3)$. (Tips: binomialsatsen $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$.) (4p)

Lösning:

$$
\begin{aligned}
P(X+Y=n) &= \sum_{k=0}^{n} P(X+Y=n, Y=k) \\
&= \sum_{k=0}^{n} P(X=n-k|Y=k)P(P=k) \\
&\overset{\text{ober.}}{=} \sum_{k=0}^{n} P(X=n-k)P(P=k) \\
&= \sum_{k=0}^{n} e^{-1}\frac{1^k}{(n-k)!}e^{-2}\frac{2^k}{k!} \\
&= e^{-1-2}\sum_{k=0}^{n} \frac{1}{n!} \cdot \frac{n!}{k!(n-k)!} \cdot 1^{n-k} \cdot 2^k \\
&= e^{-3}\frac{1}{n!}\sum_{k=0}^{n} \binom{n}{k} 1^{n-k}2^k \\
&\overset{\text{tipset}}{=} e^{-3}\frac{1}{n!}(1+2)^n \\
&= e^{-3}\frac{3^n}{n!} \\
&= P(Z=n)
\end{aligned}
$$

där $Z \in Po(3)$ dvs $X + Y \in Po(3)$. □

8. Antag att $E(X) = \mu$ och $V(X) = \sigma^2$ och att $X_1, X_2, \ldots, X_n$ är ett stickprov på X. Bevisa att

a) $\bar{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ är en väntevärdesriktig skattning av μ. (2p)

b) $\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2$ är en väntevärdesriktig skattning av σ^2. (3p)

Lösning:

a) (Se Matematisk statistik av Kerstin Vännman, s. 132.)

b) (Se Matematisk statistik av Kerstin Vännman, s. 177.)