STOCKHOLMS UNIVERSITET
MATEMATISKA INSTITUTIONEN
Avd. Matematisk statistik

MT1011
8 januari 2021

Lösningar till tentamen i Sannolikhetslära och statistik för lärare 8 januari 2021

Examinator: Jan-Olov Persson

Uppgift 1

a)
Medelvärde,

$\bar{x} = (1{\cdot}0+2{\cdot}1+1{\cdot}2+4{\cdot}3+6{\cdot}4+4{\cdot}5+1{\cdot}6+1{\cdot}7)/(1+2+1+4+6+4+1+1) = 3.65$

Median, $\tilde{x} = 4$ (storleksordna observationerna av antal utryckningar per vecka och beräkna medelvärdet av observation nummer 10 och 11).

b) För en slumpvariabel $X \sim Po(\lambda)$ är $E(X) = \lambda$ och med $\lambda = 3.65$ blir den sökta sannolikheten

$$P(X = 8) = \frac{3.65^8}{8!}e^{-3.65} = 0.0203$$

Uppgift 2

Exempel på lösning:
a) Låt P och F vara utfallet att en pojke respektive flicka föds. Utfallsrummet när två barn föds kan då skrivas som $\Omega = \{PP, PF, FP, FF\}$ där t.ex. PF innebär att det första barnet är en pojke och det andra barnet är en flicka. Eftersom det vid varje födsel är samma sannolikhet för pojke som för flicka har vart och ett av de fyra utfallen samma sannolikhet, 1/4. Sannolikheten för två flickor (FF) är alltså 1/4.

PP, PF etc. är alltså barnen *i tidsordning*.

b) Får man veta att minst ett av barnet är en flicka så kan utfallet PP uteslutas och ett reducerat utfallsrum blir $\Omega = \{PF, FP, FF\}$. Varje utfall har samma sannolikhet, 1/3 så sannolikheten för två flickor är 1/3.

c) Får man veta att det äldsta barnet är en flicka så kan både utfallet PP och PF uteslutas och ett reducerat utfallsrum blir $\Omega = \{FP, FF\}$. Varje utfall har samma sannolikhet, 1/2 så sannolikheten för två flickor är 1/2.

Uppgift 3

a) För att $P(Y = k)$ ska vara en sannolikhetsfunktion måste sannolikheterna för alla utfall summeras till 1, vilket de gör:

$$\sum_{k=0}^{k=4}(P(Y = k) = 0.2401 + 0.4116 + 0.2646 + 0.0756 + 0.0081 = 1$$

Dessutom ska $P(Y = k) \geq 0$ för alla k vilket vi ser också gäller. Slutligen får vi $P(Y \leq 2) = P(Y = 0) + P(Y = 1) + P(Y = 2) = 0.2401 + 0.4116 + 0.2646 = 0.9163$.

b) Väntevärdet av Y blir

$$E(Y) = \sum_{k=0}^{k=4} k(P(Y = k) = 0 \cdot 0.2401 + 1 \cdot 0.4116 + 2 \cdot 0.2646 + 3 \cdot 0.0756 + 4 \cdot 0.0081 = 1.2$$

Tolkning: Medelvärdet av ett stort antal oberoende utfall av Y kan förväntas bli cirka 1.2 och desto närmare 1.2 ju fler utfall som observeras.

c) En slumpvariabel $Y = \sum_{k=1}^{n} X_i$ av oberoende Bernoulli-fördelade slumpvariabler $X_i \sim Be(p)$ är binomialfördelad med väntevärde $E(Y) = np$. Vi noterar att Y som högst kan ta värdet 4 vilket innebär att antalet termer i summan måste vara precis 4, d.v.s. $n = 4$. Från uppgift b vet vi att väntevärdet av Y är 1.2 vilket ger oss ekvatoen $4p = 1.2$ med lösning $p = 0.3$. Det gäller alltså att Y är binomialfördelad med parametrar $n = 4$ och $p = 0.3$, $Y \sim Bin(n = 4, p = 0.3)$.

Uppgift 4

a)

$$P(1.2 \leq X \leq 1.4) = P(X \leq 1.4) - P(X \leq 1.2) = F(1.4) - F(1.2) = \frac{1.4^2}{4} - \frac{1.2^2}{4} = 0.13$$

b) Täthetsfunktionen blir $f(x) = F'(x) = x/2$. Med formler enligt formelbladet får man:

$$E(X) = \int_0^2 xf(x)dx = \int_0^2 x^2/2dx = \left[x^3/6\right]_0^2 = 2^3/6 = 4/3$$

$$E(X^2) = \int_0^2 x^2 f(x)dx = \int_0^2 x^3/2dx = \left[x^4/8\right]_0^2 = 2^4/8 = 2$$

$$V(X) = E(X^2) - (E(X))^2 = 2 - (4/3)^2 = 2/9 \text{ och } D(X) = \sqrt{V(X)} = \frac{\sqrt{2}}{3}$$

c) Enligt centrala gränsvärdessatsen är $\bar{X}$ approximativt normalfördelad med väntevärde $\mu = E(\bar{X}) = E(X) = 4/3$ och standardavvikelse $\sigma = D(\bar{X}) = D(X)/\sqrt{100} = \sqrt{2}/30$. Den sökta sannolikheten blir

$$P(1.2 \leq \bar{X} \leq 1.4) = \phi\left(\frac{1.4 - 4/3}{\sqrt{2}/30}\right) - \phi\left(\frac{1.2 - 4/3}{\sqrt{2}/30}\right) \approx 0.92$$

Uppgift 5

a) Låt X vara vikten på ett slumpmässigt valt äpple. Det gäller att $X \sim N(\mu = 200, \sigma = 30)$ och

$$P(200 \leq X \leq 250) = \phi\left(\frac{250 - 200}{30}\right) - \phi\left(\frac{200 - 200}{30}\right) = \phi(5/3) - \phi(0) \approx 0.45$$

b) Sannolikheten att ett slumpmässigt valt äpple väger mindre än 200 gram är $\phi(0) = 0.5$. Låt X vara antal äpplen med en vikt under 200 gram av de 5 utvalda. Då är $X \sim Bin(n = 5, p = 0.5)$ och $P(X \geq 2) = 1 - P(X \leq 1)$. Ur tabell 1 utläser man att $P(X \leq 1) = 0.1875$ vilket ger $P(X \geq 2) = 0.8125$.

c) Låt X_i vara vikten på ett slumpmässigt valt äpple och Y_j vikten på ett slumpmässigt vald apelsin för $i = 1, 2, 3, 4, 5$ och $j = 1, 2$, där $X_i \sim N(0.2, 0.03)$ och $X_j \sim N(0.3, 0.05)$. Priset för 5 äpplen och 2 apelsiner blir $V = 25\sum_{i=1}^{5} X_i + 20\sum_{j=1}^{2} Y_j$. Det gäller att $V \sim N(\mu, \sigma)$ där $\mu = E(V) = 25 \cdot 5 \cdot 0.2 + 20 \cdot 2 \cdot 0.3 = 37$ och $\sigma = D(V) = \sqrt{25^2 \cdot 5 \cdot 0.03^2 + 20^2 \cdot 2 \cdot 0.05^2} \approx 2.194$.

$$P(V > 40) = 1 - P(V \leq 40) = 1 - \phi\left(\frac{40 - 37}{2.194}\right) \approx 0.085$$

Uppgift 6

a) Ett 95% konfidensintervall för μ beräknas med formeln

$$\bar{x} \pm t_{0.025}(n-1)\frac{s}{\sqrt{n}}$$

Med $n = 5$, $\bar{x} = 1.34$, $s = 0.3435$ och $t_{0.025}(4) = 2.776$ blir intervallet 1.34 ± 0.43 eller $(0.91, 1.77)$.

b) Eftersom konfidensintervallet innesluter värdet $\mu = 1$ så kan man knappast dra slutsatsen att påståendet $\mu = 1.0$ är falskt. Mer formellt kan följande hypotestest göras: $H_0 : \mu = 1.0$ mot $H_1 : \mu \neq 1.0$. Med signifikansnivå $\alpha = 5\%$ kan nollhypotesen inte förkastas.

c) Med $n = 20$, $\bar{x} = 1.34$, $s = 0.3435$ och $t_{0.025}(19) = 2.093$ får man intervallet 1.34 ± 0.16 eller $(1.18, 1.50)$. Nu innesluter intervallet inte $\mu = 1.0$ vilket ger statistlskt stöd till slutsatsen att påståendet $\mu = 1.0$ är falskt. Med samma hypotestest som i uppgift b kan nu nollhypotesen förkastas på signifikansnivå $\alpha = 5\%$.