Dugga
MVE302/MVE395/TMA321 Sannolikhet och statistik/Sannolikhet, statistik och risk/Matematisk statistik

2021-04-27 kl. 18:00-20:00

Examinator: Johan Jonasson, Matematiska vetenskaper, Chalmers

Denna dugga utgör grund för bonuspoäng enligt information på kurshemsidan. Inlämning sker genom att ladda upp i Canvas senast kl 20:00 i pdf. Observera att samarbete inte är tillåtet och betraktas som fusk. Allt som inte innebär att ta hjälp av någon annan människa i realtid är tillåtet.

---

1. Betrakta tre urnor: urna 1 som innehåller 8 röda, 6 blå och 4 gröna bollar, urna 2 som innehåller 1 röd, 3 blå och 5 gröna bollar och urna 3 som innehåller 2 blåa och 2 gröna bollar. En av de tre urnorna väls slumpmässigt på så sätt att urna 1 väljs med sannolikhet 1/2 och urna 2 eller urna 3 med sannolikhet 1/4 vardera. Ur den valda urnan väljs sedan två bollar på måfå utan återläggning. Om du får reda på att det valdes en blå och en grön boll, vad är den betingade sannolikheten att det var urna i som valdes, $i = 1, 2, 3$?

Lösning. Låt U_i vara händelsen att urna i väljs, i=$1, 2, 3$ och låt A vara händelsen att en blå och en grön boll väljs. Då gäller

$$\mathbb{P}(U_1)\mathbb{P}(A|U_1) = \frac{1}{2} \cdot \frac{6 \cdot 4}{\binom{18}{2}} = \frac{2}{51}$$
$$\mathbb{P}(U_2)\mathbb{P}(A|U_2) = \frac{1}{4} \cdot \frac{3 \cdot 5}{\binom{9}{2}} = \frac{5}{48}$$
$$\mathbb{P}(U_3)\mathbb{P}(A|U_3) = \frac{1}{4} \cdot \frac{2 \cdot 2}{\binom{4}{2}} = \frac{1}{6}.$$

Enligt Bayes formel gäller då att

$$\mathbb{P}(U_1|A) = \frac{\mathbb{P}(U_1)\mathbb{P}(A|U_1)}{\mathbb{P}(U_1)\mathbb{P}(A|U_1) + \mathbb{P}(U_2)\mathbb{P}(A|U_2) + \mathbb{P}(U_3)\mathbb{P}(A|U_3)} = \frac{32}{253}$$

och på samma sätt $\mathbb{P}(U_2|A) = 85/253$ och $\mathbb{P}(U_2|A) = 136/253$.

2. Kasta en vanlig sexsidig tärning tills du fått två sexor. Låt X vara antalet kast du behöver göra. Bestäm frekvensfunktion, väntevärde och varians för X.

Lösning. Det lättaste sättet att lösa detta är att skriva $X = Y_1 + Y_2$ där Y_1 och Y_2 är två sv som är oberoende och $Geo(1/6)$-fördelade. Då vet vi att väntevärdena av Y_i:na är 6 och varianserna är $(5/6)/(1/6)^2 = 30$. Därför gäller

$$\mathbb{E}[X] = 12,\ \mathbb{V}\mathrm{ar}(X) = 60.$$

För att man ska få $X = k$ måste kast nr k ge en sexa och exakt ett av kasten $1, \ldots, k-1$ ska ge en sexa. Resterande $k-2$ kast ska ge något annat än en sexa. Det finns $k-1$ sätt att välja ut ett av de första $k-1$ kasten som ska ge en sexa, så frekvensfunktionen blir

$$f(k) = (k-1)\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{k-2}.$$

(Fördelningen för X kallar man för övrigt för negativ binomialfördelning med parametrar 2 och p. Allmänt säger man att X är negativt binomialfördelad med parametrar n och p om X är summan av n stycken oberoende sv som är $Geo(p)$-fördelade, vilket är ekvivalent med att säga att frekvensfunktionen är $f(k) = \binom{k-1}{n-1}p^n(1-p)^{k-n}$.)

3. Låt X vara exponentialfördelad med parameter 2 och låt Y_1 och Y_2 vara exponentialfördelade med parameter 3. Alla de tre stokastiska variablerna är oberoende. Vad är $\mathbb{P}(X > Y_1)$? Vad är $\mathbb{P}(X > Y_1 + Y_2)$?

Lösning. Enligt TSL och oberoendet mellan X och Y_1 gäller att

$$\begin{aligned}\mathbb{P}(X > Y_1) &= \int_0^\infty P(X > Y_1 | Y_1 = y) f_Y(y) dy = \int_0^\infty P(X > y) f_Y(y) dy \\ &= \int_0^\infty e^{-2y} 3e^{-3y} dy = 3 \int_0^\infty e^{-5y} dy \\ &= \frac{3}{5}.\end{aligned}$$

Här står f_Y för täthetsfunktionen för Y_1 och Y_2.

För den andra delen av uppgiften använder vi exponentialfördelningens glömskeegenskap. Först observerar vi att TSL ger

$$\mathbb{P}(X > Y_1 + Y_2) = \int_0^\infty \int_0^\infty \mathbb{P}(X > y_1 + y_2 | Y_1 = y_1, Y_2 = y_2) f_{Y_1,Y_2}(y_1, y_2) dy_1 dy_2.$$

Eftersom X är oberoende av Y_1 och Y_2 gäller

$$\mathbb{P}(X > y_1 + y_2 | Y_1 = y_1, Y_2 = y_2) = \mathbb{P}(X > y_1 + y_2).$$

Detta är i sin tur lika med $\mathbb{P}(X > y_1 + y_2 | X > y_1)\mathbb{P}(X > y_1)$ som är lika med $\mathbb{P}(X > y_1)\mathbb{P}(X > y_2)$ enligt glömskeegenskapen. Eftersom Y_1 och Y_2 är oberoende gäller också att $f_{Y_1,Y_2}(y_1, y_2) dy_1 dy_2 = f_Y(y_1) f_Y(y_2)$. Sammantaget får vi

$$\begin{aligned}\mathbb{P}(X > Y_1 + Y_2) &= \int_0^\infty \int_0^\infty \mathbb{P}(X > y_1)\mathbb{P}(X > y_2) f_Y(y_1) f_Y(y_2) dy_1 dy_2 \\ &= \int_0^\infty \mathbb{P}(X > y_1) f_Y(y_1) dy_1 \int_0^\infty \mathbb{P}(X > y_2) f_Y(y_2) dy_2 \\ &= \mathbb{P}(X > Y_1)\mathbb{P}(X > Y_2).\end{aligned}$$

Därför gäller $\mathbb{P}(X > Y_1 + Y_2) = (3/5)^2 = 9/25$.

(Det är godkänt att för den andra delen bara säga att det enligt glömskeegenskapen gäller att $\mathbb{P}(X > Y_1 + Y_2) = \mathbb{P}(X > Y_1)\mathbb{P}(X > Y_2)$. Det påståendet är dock inte allmänt sant om inte Y_1 och Y_2 är oberoende, så detta ska man egentligen också anföra som skäl. Det gäller till exempel att $\mathbb{P}(X > 2Y_1) = 3/7 \neq \mathbb{P}(X > Y_1)^2$.)

Lycka till!
Johan Jonasson