Aalto-universitetet

MS-A0509 Grundkurs i sannolikhetskalkyl och statistik — Gripenberg
Övning 2
Vecka 3, 12–16.1.2015

I1. En käpp med längden 6 (meter) delas i två delar så att den ena delen har längden X och den andra längden $6-X$ där X är en slumpvariabel som är jämnt fördelad i intervallet $[0,6]$. Båda delarna delas i mitten och av delarna bildas en rektangel med sidoren $\frac{1}{2}X$ och $\frac{1}{2}(6-X)$. Bestäm väntevärdet av arean av rektangeln. Är det en god idé att uppskatta väntevärdet genom att ta väntevärdet av X som värdet på X?
Ledning: Använd dig endast av täthetsfunktionen för X och formeln för väntevärdet av en funktion av en slumpvariabel och räkna alltså inte ut täthetsfunktionen eller fördelningsfunktionen för arean!

Svar: $\frac{3}{2}$

Lösning: Arean av rektangeln är $\frac{1}{2}X\cdot\frac{1}{2}(6-X)=\frac{1}{4}X(6-X)$ och vi skall alltså räkna ut $\mathrm{E}(\frac{1}{4}X(6-X))$. Eftersom X är är jämnt fördelad i intervallet $[0,6]$ är dess täthetsfunktion $f_X(t)=\frac{1}{6}$ då $0\le t\le 5$ och 0 annars. Väntevärdet blir därför

$$\begin{aligned}\mathrm{E}\left(\tfrac{1}{4}X(6-X)\right)&=\frac{1}{4}\int_0^6 t(6-t)\frac{1}{6}\,\mathrm{d}t\\ &=\frac{1}{24}\int_0^6\left(6t-t^2\right)\,\mathrm{d}t=\frac{1}{24}\Big/_0^6\left(3t^2-\frac{1}{3}t^3\right)=\frac{1}{24}\left(3\cdot 36-\frac{1}{3}\cdot 6\cdot 36\right)=\frac{1}{24}\cdot 36=\frac{3}{2}.\end{aligned}$$

Det är inte en god idé att uppskatta väntevärdet genom att ta väntevärdet av X som värdet på X eftersom väntevärdet av X som är 3 ger maximivärdet av arean.

I2. Du reser på solsemester till Ologa, där sannolikheten att solen skiner är p och du har bestämt att du far hem genast när du har tillbringat $r\ge 1$ soliga dagar där. Använd följande resonemang för att härleda ett uttryck för sannolikheten att du stannar n dagar i Ologa om du antar att händelserna *"Dagen j är solig"* är oberoende för alla j (och denna sannolikhet är alltså värdet av frekvensfunktionen för den negativa binomialfördelningnen i punkten n med parametrarna p och r dvs. $f_{\mathrm{NegBin}(p,r)}(n)$):

(a) Låt A vara händelsen *"Du tillbringar n dagar i Ologa"*, B händelsen *"Dag nummer n är solig"* och C händelsen *"Av de $n-1$ första dagarna är $r-1$ soliga"*. En av kombintionerna av händelser A och B, A och C, eller B och C är säkert oberoende? Vilken är det?
(b) Uttryck händelsen A med hjälp av händelserna B och C.
(c) Bestäm $\Pr(B)$ och $\Pr(C)$.
(d) Bestäm $\Pr(A)$.

Lösning: (a) Händelserna B och C är oberoende eftersom vi antar att händelserna *"Dagen j är solig"* är oberoende men varken händelserna A och C eller A och B är oberoende.

(b) Om dag nummer n är solig och av de föregånde $n-1$ dagarna $r-1$ dagar har varit soliga så har du efter dag nummer n upplevt $r-1+1=r$ soliga dagar och reser hem, dvs. $A=B\cap C$.

(c) Enligt antagandet är $\Pr(B) = p$. Eftersom händelserna *"Dagen j är solig"* är oberoende och har sannolikheten p för alla j så får vi $\Pr(C)$ med hjälp av binmoialfördelningen och $\Pr(C) = \binom{n-1}{r-1} p^{r-1}(1-p)^{n-1-r+1}$.

(d) Eftersom B och C är oberoende och $A = B \cap C$ så gäller

$$\Pr(A) = \Pr(B \cap C) = \Pr(B)\Pr(C) = p\binom{n-1}{r-1} p^{r-1}(1-p)^{n-1-r+1} p = \binom{n-1}{r-1} p^r (1-p)^{n-r}.$$

I3. En kommitté har 7 kvinnliga och 6 manlig medlemmar. Av dessa väljs slummässigt medlemmar till en arbetsgrupp, en i taget, tills det finns 3 kvinnor i arbetsgruppen. Vad är sannolikheten att det då finns 5 personer i arbetsgruppen, 3 kvinnor och 2 män?
Ledning: Räkna först ut sannolikheten att då det valts 4 medlemmar till arbetsgruppen så består den av 2 kvinnor och 2 män.

Svar: $\frac{35}{143}$

Lösning: Om slutresultatet är att det finns 5 personer i arbetsgruppen så måste det ha valts 2 kvinnor och 2 män då det valts 4 personer eftersom den sista som väljs är en kvinna. Om alltså

$$\begin{aligned}
A &= \{\text{"Arbetsgruppen består av 5 personer"}\} \\
B &= \{\text{"Av de 4 första valda är 2 kvinnor och 2 män"}\} \\
C &= \{\text{"Den femte som väljs är en kvinna"}\}.
\end{aligned}$$

så är $A = B \cap C$ så att vi får

$$\Pr(A) = \Pr(B \cap C) = \Pr(C|B)\Pr(B).$$

Nu är

$$\Pr(B) = \frac{|\text{"Välj 2 kvinnor bland 7"}| \cdot |\text{"Välj 2 män bland 6"}|}{|\text{"Välj 4 personer bland 13"}|} = \frac{\binom{7}{2} \cdot \binom{6}{2}}{\binom{13}{4}}.$$

Om vi först har valt 2 kvinnor och 2 män så finns det kvar 5 kvinnor och 4 män och då är sannolikheten för att den nästa som väljs är en kvinna

$$\Pr(C|B) = \frac{5}{9},$$

och vi får

$$\Pr(A) = \frac{\binom{7}{2} \cdot \binom{6}{2}}{\binom{13}{4}} \cdot \frac{5}{9} = \frac{35}{143}.$$

I4. Antag att livslängden för en viss komponent i en maskin är exponentialfördelad och antag också att 90% av dessa komponenter inte fungerar efter 60 dagar. Hur många gånger måste man i genomsnitt byta ut denna komponent mot en ny under ett års tid (om man alltså byter den genast då den gått sönder), dvs. vad är väntevärdet av antalet byten?

Ledning: Räkna först ut parametern för exponentialfördelningen genom att använda uttrycket för fördelningsfunktionen och kom sedan ihåg sambandet mellan exponential- och Poissonfördelningen (vilket säger att sambandet mellan den genomsnittliga livslängden och antal byten är enklast tänkbara just för exponentialfördelningen).

Svar: ≈ 14

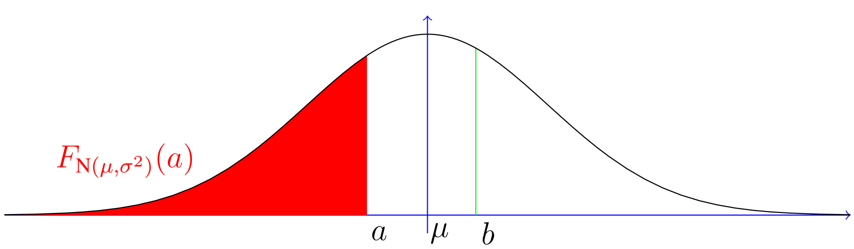
Lösning: Om $X \sim \mathrm{Exp}(\lambda)$ så är dess fördelningsfunktion $F_X(t) = \Pr(X \leq t) = 1 - \mathrm{e}^{-\lambda t}$. Eftersom $\Pr(X \leq 60) = 0.9$ så är

$$0.9 = 1 - \mathrm{e}^{-\lambda \cdot 60} \quad \Rightarrow \lambda = -\frac{\log(0.1)}{60} = 0.038376.$$

Om nu X_j är oberoende Expλ)-fördelade slumpvariabler, $T > 0$ och $Y = \max\{\, k \geq 0 : X_1 + X_2 + \ldots X_k \leq T \,\}$ så gäller $Y \sim$ Poisson(λT) och E$(Y) = \lambda T$. Om $T = 365$ så är Y antalet byten som måste göras under ett år och väntevärdet av Y är $-\frac{\log(0.1)}{60} \cdot 365 = 14.007$.

I5. Antag att vikten av fiskar som fångats är N(μ, σ^2)-fördelad. Alla fiskar vars vikt är högst a ges åt kattorna. Bestäm medianen av viktfördelningen av de fiskar som inte ges åt kattorna och ge ditt svar med hjälp av parametrarna μ, σ, a och N$(0,1)$-fördelningens fördelningsfunktion $F_{\mathrm{N}(0,1)}$ och dess inversa funktion.

Ledning: Hälften av de fiskar vars vikt är större än a väger mindre än den median som skall bestämmas och hälften väger mera. Kom också ihåg att $F_{\mathrm{N}(\mu,\sigma^2)}(t) = F_{\mathrm{N}(0,1)}(\frac{t-\mu}{\sigma})$ och att arean under täthetsfunktionen i ett intervall är sannolikheten för att slumpvariablen får sitt värde i detta intervall.



Lösning: Om X är vikten av en fisk så gäller enligt antagandet $X \sim \mathrm{N}(\mu, \sigma^2)$ så att $\dfrac{X-\mu}{\sigma} \sim$ N$(0,1)$.

Om medianen av vikten av de fiskar som väger mera än a är b så följer det av definitionen av medianen (och av det faktum att X är en kontinuerlig slumpvariabel) att

$$\Pr(a < X < b) = \Pr(X > b).$$

Av detta följer att

$$\Pr(X > b) = \frac{1}{2}\Pr(X > a) = \frac{1}{2}(1 - \Pr(X \leq a)),$$

så att

$$\Pr(X \leq b) = 1 - \frac{1}{2}(1 - \Pr(X \leq a)) = \frac{1}{2}(1 + \Pr(X \leq a)).$$

Eftersom $X \leq b$ om och endast om $\dfrac{X-\mu}{\sigma} \leq \dfrac{b-\mu}{\sigma}$ så gäller

$$\frac{b-\mu}{\sigma} = F^{-1}_{\mathrm{N}(0,1)}\left(\frac{1}{2}\left(1 + F_{\mathrm{N}(0,1)}\left(\frac{a-\mu}{\sigma}\right)\right)\right),$$

och medianen är

$$\mu + \sigma F_{\mathrm{N}(0,1)}^{-1}\left(\frac{1}{2}\left(1 + F_{\mathrm{N}(0,1)}\left(\frac{a-\mu}{\sigma}\right)\right)\right).$$