Solution to Exercise C−1

At the time of the k-th draw (if the game continues this long), there are $k+1$ dollars in the pot, one of them marked, the other k unmarked. Thus

$$\begin{aligned}
&P(Y=k) \\
&= P(\text{first } k-1 \text{ draws are } \textbf{unmarked}\text{, and the } k\text{th draw is marked}) \\
&= \frac{1}{2} \times \frac{2}{3} \times \cdots \times \frac{k-1}{k} \times \frac{1}{k+1} \\
&= \frac{1}{k(k+1)}
\end{aligned}$$

for $k = 1, 2, 3, \ldots$. This pmf was discussed in lecture. Recall that (since Y is nonnegative)

$$E(|Y|) = EY = \sum_{k=1}^{\infty} k \cdot \frac{1}{k(k+1)} = \sum_{k=1}^{\infty} \frac{1}{(k+1)} = \infty.$$

Thus EY is undefined (or we could safely regard it as $+\infty$ in this case). The draws proceed A, B, A, B, $\ldots$ so that A makes all the odd draws, and B makes all the even draws. If $Y = k$ and k is odd, then A wins $(k+1)/2$ dollars. (There are $k+1$ dollars in the hat when A wins, but he put half of them in himself.) If $Y = k$ and k is even, then A loses $(k+2)/2$ dollars. Thus A's winnings can be expressed as a function of the length of the game

$$Z = g(Y) = \begin{cases} (Y+1)/2 & \text{if } Y \text{ is odd} \\ -(Y+2)/2 & \text{if } Y \text{ is even} \end{cases}$$

so that

$$\begin{aligned}
E(|Z|) &= E(|g(Y)|) = \sum_{k=1}^{\infty} |g(k)| f_Y(k) \\
&= \sum_{k \text{ odd}} \frac{k+1}{2} \cdot \frac{1}{k(k+1)} + \sum_{k \text{ even}} \frac{k+2}{2} \cdot \frac{1}{k(k+1)} \\
&\quad \text{(Now replace } k+2 \text{ by the smaller value } k+1 \text{ in the} \\
&\quad \text{second sum and cancel factors of } k+1.) \\
&\geq \sum_{k=1}^{\infty} \frac{1}{2k} = \infty\,.
\end{aligned}$$

Thus EZ is undefined. (In this case one cannot safely regard EZ as being either $+\infty$ or $-\infty$; it is simply undefined.)

Discussion: Since EZ is undefined, if this game is played repeatedly, the Law of Large Numbers fails, and there is no such thing as a "stable long run average" for A's winnings. Suppose A and B play this game a million times (assume A and B are immortal, infinitely wealthy, and have nothing better to do) and let $\bar{Z}_1$ be "A's average winnings per game" over this series of a million games. Now suppose A and B play another million games, and let $\bar{Z}_2$ be A's average winnings for this second series. The two values $\bar{Z}_1$ and $\bar{Z}_2$ will **not** (except by coincidence) be close to each other. If we replace "million" by "billion" or "trillion", the two values will still not be close (except by coincidence).

Problem CO

$$P(A \cup B \cup C) = 1 - P((A \cup B \cup C)^c)$$

$$= 1 - P(A^c \cap B^c \cap C^c)$$

$$= 1 - E(I_{A^c \cap B^c \cap C^c}). \qquad (*)$$

Now note that

$$I_{A^c \cap B^c \cap C^c} = I_{A^c} \cdot I_{B^c} \cdot I_{C^c}$$

$$= (1 - I_A)(1 - I_B)(1 - I_C)$$

$$= 1 - I_A - I_B - I_C + I_A \cdot I_B + I_A \cdot I_C$$
$$+ I_B \cdot I_C - I_A \cdot I_B \cdot I_C$$

$$= 1 - I_A - I_B - I_C + I_{A \cap B} + I_{A \cap C}$$
$$+ I_{B \cap C} - I_{A \cap B \cap C}$$

Plugging this into (*) and taking expectations gives

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C)$$
$$- P(B \cap C) + P(A \cap B \cap C).$$

Heather Schonrock
C1

Find the mean and variance of the player's total winnings.

Let X_i = i^{th}, $i+1^{th}$, and $i+2^{th}$ are correct

mean = $E\sum_{i=1}^{10} X_i = \sum_{i=1}^{10} EX_i$

EX_i for any X_i is $0(7/8)+1(1/8)=\frac{1}{8}$

$\therefore \sum_{i=1}^{10} EX_i = \sum_{i=1}^{10} 1/8 = \boxed{10/8}$

let $S_K = \sum_{i=1}^{10} X_i$

variance = $ES_K^2 - (ES_K)^2$

$\qquad\qquad\qquad\quad \overset{\shortparallel}{(10/8)^2}$

$E(S_K^2) = \sum_{i=1}^{10} EX_i + 2\sum_{i<j}\sum EX_iX_j$ (done in class)

$= \frac{10}{8} + \frac{160}{64}$ (see below)

To work out $2\sum\sum EX_iX_j$, all the possibilities can be written out. Note that not all cases are disjoint. For example, $EX_1X_2 = 1/16$ because X_1 is 1,2,3 correct and X_2 is 2,3,4 correct.

$\therefore 2\sum_{i<j}\sum EX_iX_j = 2\left(\frac{80}{64}\right) = \frac{160}{64}$

$\therefore$ Variance $= \frac{10}{8} + \frac{160}{64} - \left(\frac{10}{8}\right)^2 = \frac{140}{64} = 2\frac{3}{16}$

[C1](Supplement to Solution)

$$X_i = I_{\{i^{th}, (i+1)^{th}, (i+2)^{th} \text{ correct}\}}$$

$$S = \text{winnings} = \sum_{i=1}^{10} X_i$$

$$ES^2 = \sum_{i=1}^{10} \sum_{j=1}^{10} E X_i X_j$$ (100 terms. Arrange in a 10x10 matrix.)

- 10 terms with $i=j$:

$$E X_i^2 = E X_i = P(i^{th}, (i+1)^{th}, (i+2)^{th} \text{ correct})$$
$$= \left(\tfrac{1}{2}\right)^3 = \tfrac{1}{8}$$

- $9 \times 2 = 18$ terms with $|i-j| = 1$:

$$E X_i X_{i+1} = P(i, i+1, i+2, i+3 \text{ correct})$$
$$= \left(\tfrac{1}{2}\right)^4 = \tfrac{1}{16}$$

- $8 \times 2 = 16$ terms with $|i-j| = 2$:

$$E X_i X_{i+2} = P(i, i+1, i+2, i+3, i+4 \text{ correct})$$
$$= \left(\tfrac{1}{2}\right)^5 = \tfrac{1}{32}$$

- $100 - 10 - 18 - 16 = 56$ terms with $|i-j| \geq 3$ (no overlap)

$$E X_i X_j = P(i, i+1, i+2, j, j+1, j+2 \text{ correct})$$
$$= \left(\tfrac{1}{2}\right)^6 = \tfrac{1}{64}.$$

Thus

$$ES^2 = 10 \cdot \frac{1}{8} + 18 \cdot \frac{1}{16} + 16 \cdot \frac{1}{32} + 56 \cdot \frac{1}{64}$$

$$= (80 + 72 + 32 + 56)/64$$

$$= 240/64 = 15/4$$

$$\mathrm{Var}\, S = ES^2 - (ES)^2 = \frac{15}{4} - \left(\frac{10}{8}\right)^2$$

$$= \frac{35}{16} = 2\frac{3}{16}$$

Inference I

Ivo Dinov

Problem [C.2] Consider a nation of $N = 2\times 10^8$ inhabitants. the set of all people Ω, splits as a disjoint union: $\Omega = A \cup B \cup C$, with
$P(\text{someone from A run amok}) = P_A = \frac{1}{10^9}$
$P(\text{someone from B run amok}) = P_B = \frac{1}{10^8}$
$P(\text{someone from C run amok}) = P_C = \frac{1}{10^7}$, and
$|A| + |B| + |C| = N$, $|A| = 15\times 10^7$, $|B| = 4\times 10^7$,
$|C| = 10^7$.
Find an approximation to the probability that exactly 2 people run amok.

Solution: Let X_i be the random variable $= \begin{cases} 1, & \text{if } i^{th} \text{ person runs amok} \\ 0, & \text{otherwise} \end{cases}$

since people act independently $\Rightarrow$
$\{X_i\}_{i=1}^N$ are independent, with $X_i \sim B(n_\alpha, P_\alpha)$ for $\alpha = A, B, C$.
Thus, since the total # of $\{X_i\}_1^N$ is N-large and all P_A, P_B, P_C are small

$$\Rightarrow X := \sum_1^N X_i \approx Y, \text{ with } Y \sim \text{Poisson}\Big(\underbrace{\sum_{i=1}^N p_i}_{\lambda}\Big),$$

where $p_i = P(i^{th} \text{ person runs amok})$.

$$\Rightarrow P(X=2) \approx P(Y=2) = \frac{\lambda^2 e^{-\lambda}}{2},$$

$$\lambda = \sum_1^N p_i = \sum_A P_A + \sum_B P_B + \sum_C P_C = \frac{1}{10^9}15\times 10^7 + \frac{4\times 10^7}{10^8} + \frac{10^7}{10^7}$$

$$\Rightarrow \lambda = \frac{15}{100} + \frac{40}{100} + \frac{100}{100} = \frac{155}{100}$$

$$\Rightarrow P(X=2) \approx 0.255$$

$\boxtimes$

Clay Venable

C-3 Let $X_1, \ldots, X_{10}$ be iid RV's with density $f(x) = 2x \quad x \in (0,1)$

Let $Y = \max X_i$ and $Z = \min X_i$

ⓐ Find cdf of Y and EY

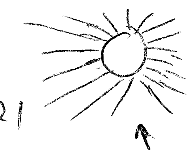
$$P(Y \le y) = P(X_1 \le y \text{ and } \ldots X_{10} \le y) = P(X_1 \le y)P(X_2 \le y) \cdots P(X_{10} \le y)$$

THIS CALLS FOR THE POWER OF... CALCULUS MAN!!

$$= \left(\int_0^y dt\, 2t\right)^{10} = y^{20}$$

$$y^{20} = \int_{-\infty}^{y} f_Y(t)\,dt = \int_0^y f_Y(t)\,dt$$

so $f_Y(y) = 20y^{19}$ $\longrightarrow$ note: $\int_0^1 20y^{19}\,dy = y^{20}\Big|_0^1 = 1$



$$EY = \int_{\mathbb{R}} y\, 20y^{19} = \int_0^1 20y^{20} = \frac{20}{21} y^{21}\Big|_0^1 = \frac{20}{21}$$

SUN SHINING ON EXPECTED VALUE

ⓑ Find cdf of Z and EZ OK!

$$f_Z(z) = P(Z = z) = P(Z_1 = z \text{ and } Z_1 \le Z_2, \ldots, Z_{10}) + P(Z_2 = z \text{ and } Z_2 \le Z_1, Z_3, Z_4 \ldots Z_{10}) \ldots$$

$$= 10\,P(Z_i = z)(1 - F_{Z_i}(z))^9 \longrightarrow \text{note } F_{Z_i}(z) = \int_0^z 2t\,dt = z^2$$

$$= 20z(1-z^2)^9$$

$$F_Z(z) = \int_{-\infty}^{z} 20t(1-t^2)^9\,dt = \int_0^z 20t(1-t^2)^9\,dt = -(1-t^2)^{10}\Big|_0^z = 1-(1-z^2)^{10}$$

$$EZ = \int_{\mathbb{R}} z\, f_Z(z)\,dz = \int_0^1 z\, 20z(1-z^2)^9\,dz = -\int_0^1 z^2\, 20(z^2-1)^9\,dz$$

$$= -\int_0^1 20z^2\left(z^{18} - 9z^{16} + 36z^{14} - 84z^{12} + 126z^{10} - 126z^8 + 84z^6 - 36z^4 + 9z^2 - 1\right)dz$$

WHEW!

$$= -20\int_0^1 z^{20} - 9z^{18} + 36z^{16} - 84z^{14} + 126z^{12} - 126z^{10} + 84z^8 - 36z^6 + 9z^4 - z^2\,dz$$

$$\approx .2702602$$

[C3] (Supplement: Computational Shortcut)

cdf of Z was found to be

$$F_Z(z) = \underbrace{1-(1-F(z))^n}_{\text{general formula}} = 1-(1-z^2)^{10}, \quad 0<z<1.$$

pdf:

$$f_Z(z) = F_Z'(z) = 20z(1-z^2)^9, \quad 0<z<1$$

$$EZ = \int_0^1 z f_Z(z)\,dz = \int_0^1 20z^2(1-z^2)^9\,dz$$

(Let $u=z^2$, $z=\sqrt{u}$, $du = 2z\,dz$)

$$= 10\int_0^1 \sqrt{u}\,(1-u)^9\,du$$

$$= 10\int_0^1 u^{3/2-1}(1-u)^{10-1}\,du$$ (Now use "Beta integral". See Beta pdf.)

$$= 10\,\frac{\Gamma(\frac{3}{2})\Gamma(10)}{\Gamma(\frac{3}{2}+10)} = \frac{10\cdot 9!}{\frac{21}{2}\cdot\frac{19}{2}\cdot\dots\cdot\frac{3}{2}}$$

$$= \frac{10!\,2^{10}}{21\cdot 19\cdot 17\cdot\dots\cdot 3} = .2702602$$

Sonya Bowens

Problem C4

Three lousy light bulbs with independent lifetimes which have an exponential distribution with a mean of 4 hours. When working under 3 light bulbs Bob makes errors at an avg. of 1 per hour; 2 bulbs makes errors at an avg. of 3 per hour; 1 bulb makes errors at an avg of 6 per hour. What is expected value of the total number of errors that Bob makes?

Let Z_1 = time it takes for 1st light bulb to fail
" Z_2 = " " " " 2nd " " " "
" Z_3 = " " " " 3rd " " " "

Let λ_1 = rate of errors made until 1st light bulb fails
= 1 per hour

Let λ_2 = rate of errors made after 1st light bulb fails
= 3 per hour

Let λ_3 = rate of errors made after 2nd light bulb fails
= 6 per hour

The expected value of the total # of errors is the sum of the time it takes for the light bulbs to go out multiplied by the error rate.

$$E(\text{Total Errors}) = Z_1(\lambda_1) + Z_{22}(\lambda_2) + Z_{33}(\lambda_3)$$

where Z_{22} is the time it takes for the second light bulb to go out minus the 1st $(Z_2 - Z_1)$

Similarly for $Z_{33} = (Z_3 - Z_2)$

From Class we derived that for an exponential dist^n the failure rate = $1/\text{mean}$. Therefore $EZ_1 = \beta/3$, $E(Z_{22}) = \beta/2$, $E(Z_{33}) = \beta$

E(Total Errors) = is the sum expected value of the separate failures multiplied by the respective error rate.

$$E(\text{Total Errors}) = EZ_1(\lambda_1) + EZ_{22}(\lambda_2) + EZ_{33}(\lambda_3)$$

$$= \beta/3\,(1) + \beta/2\,(3) + \beta\,\cancel{z}(6)$$

$$= 4/3 + 4/2(3) + 4(6)$$

$$= 4/3 + 12/2 + 24$$

$$= 31\tfrac{1}{3}$$

Yoonjung Lee.

C5. Let X = # of people voting for A.

Let Y = # of people voting for B.

Then $X \sim B(2{,}000{,}000, 0.8)$

and $Y \sim B(2{,}000{,}000, 0.8)$.

So $X \dot\sim N(1{,}600{,}000, 320{,}000)$.

and $Y \dot\sim N(1{,}600{,}000, 320{,}000)$.

Assume Joe supports B.

Let $D = X - Y$.

Then $D \dot\sim N(0, 640000)$.

Joe breaks a tie.

Joe creates a tie.

$p(\text{Joe makes difference}) = p(0 \leq D \leq 1)$

$$\doteq p\left(\frac{0-0.5-0}{800} \leq Z \leq \frac{1+0.5-0}{800}\right)$$

(by continuity correction).

$$= p(-.000625 \leq Z \leq 0.001875)$$

$$= \int_{-.000625}^{0.001875} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

$\doteq .000997356$ (by using the calculator).

$$\approx \underbrace{\frac{1}{\sqrt{2\pi}}}_{\text{density at } z=0} \underbrace{(.001875-(-.000625))}_{\text{length of interval of integration}}$$

MSims

Exercises C

#6 (a)

$X \overset{d}{\approx} \text{Beta}(30, 30)$

$EX = \frac{30}{30+30} = .5$

$\text{Var}X = \frac{30(30)}{(30+30)^2(30+30+1)} \approx .0041$

$Z = \frac{X - EX}{\sigma} = \frac{X - .5}{.064} \overset{d}{\approx} N(0,1)$

$P(.4 < X < .6)$

$= P\left(\frac{.4-.5}{.064} < Z < \frac{.6-.5}{.064}\right)$

$= P(-1.5625 < Z < 1.5625)$

$= 1 - 2P(Z > 1.5625)$

$= 1 - 2(.0591)$

$= .8818$

Using pbeta in Splus, you get .880896.

Exercises C

6 (b)

$X \sim \text{Poisson}(80)$

$EX = 80$

$\text{Var}\, X = 80$

$$Z = \frac{X-80}{\sqrt{80}} \sim N(0,1)$$

$$P(75 \le X \le 90) \approx P(74.5 \le X \le 90.5)$$

$$= P\left(\frac{74.5-80}{\sqrt{80}} \le Z \le \frac{90.5-80}{\sqrt{80}}\right)$$

$$= P(-.615 \le Z \le 1.174)$$

$$= P(Z \le 1.174) - P(Z \le -.615)$$

$$= 1 - .1202 - .26925$$

$$= .61055$$

Using ppois in Splus, you get .60541

Solution to Exercise C7

Let $X \sim \text{Geometric}(p)$. Recall that

$$EX = \frac{1}{p} \quad \text{and} \quad EX^2 = \frac{2-p}{p^2}\,.$$

$$\begin{aligned}
EX^3 &= E(X^3 \,|\, X=1)P(X=1) + E(X^3 \,|\, X>1)P(X>1) \\
&= 1\cdot p + E\left\{(1+(X-1))^3 \,|\, X>1\right\}\cdot(1-p) \\
&= 1\cdot p + E\left\{(1+X)^3\right\}\cdot(1-p) \\
&= 1\cdot p + E\left(1+3X+3X^2+X^3\right)\cdot(1-p) \\
&= p + \left(1+3EX+3EX^2+EX^3\right)(1-p) \\
&= p + \left(1+\frac{3}{p}+\frac{3(2-p)}{p^2}+EX^3\right)(1-p) \\
&= \frac{p^2-6p+6}{p^2} + (1-p)EX^3
\end{aligned}$$

so that

$$pEX^3 = \frac{p^2-6p+6}{p^2} \quad \text{and} \quad EX^3 = \frac{p^2-6p+6}{p^3}\,.$$

Note: The pmf of the Geometric distribution is a geometric series, and if you remove the first few terms from a geometric series, what remains is another geometric series. This observation is essentially equivalent to the Discrete Memoryless Property. The above calculation can be rewritten as

$$\begin{aligned}
EX^3 &= \sum_{x=1}^{\infty} x^3(1-p)^{x-1}p = p + \sum_{x=2}^{\infty} x^3(1-p)^{x-1}p \\
&= p + (1-p)\sum_{x=2}^{\infty}\left((x-1)+1\right)^3(1-p)^{(x-1)-1}p
\end{aligned}$$

so that making the change of variable $z = x-1$ leads to

$$\begin{aligned}
&= p + (1-p)\sum_{z=1}^{\infty}(z+1)^3(1-p)^{z-1}p = p + (1-p)E\left[(X+1)^3\right] \\
&= p + (1-p)\left(1+3EX+3EX^2+EX^3\right) \quad \text{same as before.}
\end{aligned}$$