ECE 4110/5110 Friday, 08/25/23

Homework Set 1 Solution

Dr. Kevin Tang HW 1

1 Problems from Bertsekas and Tsitsiklis, 2nd edition

Problem 1 (Ch2 P32) Consider 2m persons forming m couples who live together at a given time. Suppose that at some later time, the probability of each person being alive is p, independent of other persons. At that later time. let A be the number of persons that are alive and let S be the number of couples in which both partners are alive. For any survivor number a, find $E[S|A = a]$.

Solution

Let X_i be the random variable taking the value 1 or 0 depending on whether the first partner of the ith couple has survived or not. Let Y_i be the corresponding random variable for the second partner of the ith couple. Then, we have $S = \sum_{i=1}^{m} X_i Y_i$, and by using the total expectation theorem,

$$\begin{aligned}
E[S|A = a] &= \sum_{i=1}^{m} E[X_i Y_i | A = a] \\
&= mE[X_1 Y_1 | A = a] \\
&= mE[Y_1 | X_1 = 1, A = a]P(X_1 = 1|A = a) \\
&= mP(Y_1 = 1|X_1 = 1, A = a)P(X_1 = 1|A = a).
\end{aligned}$$

We have

$$P(Y_1 = 1|X_1 = 1, A = a) = \frac{a-1}{2m-1}, \qquad P(X1 = 1|A = a) = \frac{a}{2m}$$

Thus

$$E[S|A = a] = \frac{a(a-1)}{2(2m-1)}$$

Note that $E[S|A = a]$ does not depend on p.

Problem 2 (Ch3 P20) An absent-minded professor schedules two student appointments for the same time. The appointment durations are independent and exponentially distributed with mean thirty minutes. The first student arrives on time, but the second student arrives five minutes late. What is the expected value of the time between the arrival of the first student and the departure of the second student?

Solution

Let T_1, T_2 be the staying times of the two students. Then we can write the overall time in question as $T = \max(T_1, 5) + T_2$ with its expectation as

$$\begin{aligned} E[T] &= E[\max(T_1, 5) + T_2] \\ &= E[\max(T_1, 5)] + E[T_2] \\ &= 5 \cdot P(T_1 \leq 5) + E[T_1 | T_1 > 5] \cdot P(T_1 > 5) + E[T_2] \end{aligned}$$

We have $E[T_2] = 30$, and using the memoryless property of exponential distribution,

$$E[T_1 | T_1 > 5] = 5 + E[T_1] = 35.$$

Also

$$\begin{aligned} P(T_1 \leq 5) = 1 - e^{-5/30}, \\ P(T_1 > 5) = e^{-5/30} \end{aligned}$$

By substitution we obtain

$$E[T] = 5 \cdot (1 - e^{-5/30}) + 35 \cdot e^{-5/30} + 30 = 60.394$$

Problem 3 (Ch3 P25) The coordinates X and Y of a point are independent zero mean normal random variables with common variance σ^2. Given that the point is at a distance of at least c from the origin, find the conditional joint PDF of X and Y.

Solution
Let C denote the event that $X^2+Y^2 \geq c^2$. The probability $P(C)$ can be calculated using polar coordinates, as follows:

$$\begin{aligned} P(C) &= \frac{1}{2\pi\sigma^2}\int_0^{2\pi}\int_c^{\infty} re^{-r^2/2\sigma^2} dr d\theta \\ &= \frac{1}{\sigma^2}\int_c^{\infty} re^{-r^2/2\sigma^2} dr \\ &= e^{-c^2/2\sigma^2} \end{aligned}$$

Thus, for $(x, y) \in C$,

$$f_{X,Y|C}(x, y) = \frac{f_{X,Y}(x, y)}{P(C)} = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(x^2+y^2-c^2)}$$

Problem 4 (Ch4 P18) Consider four random variables, W, X, Y, Z with

$$E[W] = E[X] = E[Y] = E[Z] = 0,$$

$$var(W) = var(X) = var(Y) = var(Z) = 1.$$

and assume that W. X, Y. Z are pairwise uncorrelated. Find the correlation coefficients $\rho(R, S)$ and $\rho(R, T)$. where R = W + X, S = X + Y. and T = Y + Z.

Solution

$$cov(R, S) = E[RS] - E[R]E[S] = E[WX + WY + X^2 + XY] = E[X^2] = 1$$

and

$$var(R) = var(S) = 2$$

so

$$\rho(R, S) = \frac{cov(R, S)}{\sqrt{var(R)var(S)}} = \frac{1}{2}$$

We also have

$$cov(R, T) = E[RT] - E[R]E[T] = E[WY + WZ + XY + XZ] = 0$$

so that

$$\rho(R, T) = 0$$

Problem 5 (Ch4 P19) Suppose that a random variable X satisfies

$$E[X] = 0, \quad E[X^2] = 1, \quad E[X^3] = 0, \quad E[X^4] = 3$$

and let $Y = a + bX + cX^2$. Find the correlation coefficient p(X, Y).

Solution

To compute the correlation coefficient

$$\rho(X, Y) = \frac{cov(X, Y)}{\sigma_X \sigma_Y}$$

we first compute the covariance

$$\begin{aligned} cov(X, Y) &= E[XY] - E[X]E[Y] \\ &= E[aX + bX^2 + cX^3] - E[X]E[Y] \\ &= aE[X] + bE[X^2] + cE[X^3] \\ &= b \end{aligned}$$

We also have

$$\begin{aligned} var(Y) &= var(a + bX + cX^2) \\ &= E[(a + bX + cX^2)^2] - (E[a + bX + cX^2])^2 \\ &= (a^2 + 2ac + b^2 + 3c^2) - (a^2 + c^2 + 2ac) \\ &= b^2 + 2c^2 \end{aligned}$$

Finally

$$\rho(X, Y) = \frac{b}{\sqrt{b^2 + 2c^2}}$$

Problem 6 (Ch4 P24) A retired professor comes to the office at a time that is uniformly distributed between 9 a.m. and 1 p.m., performs a single task and leaves when the task is completed. The duration of the task is exponentially distributed with parameter, $\lambda(y) = 1/(5-y)$, where y is the length of the time interval between 9 a.m. and the time of his arrival:

(a) What is the expected amount of time that the professor devotes to the task?

(b) What is the expected time at which the task is completed?

(c) The professor has a Ph.D. student who on a given day comes to see him at a time that is uniformly distributed between 9 a.m. and 5 p.m. If the student does not find the professor, he leaves and does not return. If he finds the professor, he spends an amount of time that is uniformly distributed between 0 and 1 hour. The professor will spend the same total amount of time on his task regardless of whether he is interrupted by the student. What is the expected amount of time that the professor will spend with the student and what is the expected time at which he will leave his office?

Solution

a) Consider the following two random variables:

X = amount of time the professor devotes to his task [exponentially distributed with parameter $\lambda(y) = 1/(5-y)$]
Y = length of time between 9 a.m. and his arrival (uniformly distributed between 0 and 4).

Note that $E[Y] = 2$. We have

$$E[X|Y = y] = \frac{1}{\lambda(y)} = 5 - y$$

which implies that

$$E[X|Y] = 5 - Y$$

and

$$E[X] = E\left[E[X|Y]\right] = E[5 - Y] = 5 - E[Y] = 5 - 2 = 3$$

b) Let Z be the length of time 9 a.m. until the professor completes the task. Then,

$$Z = X + Y$$

We already know from part (a) that $E[X] = 3$ and $E[Y] = 2$ so that

$$E[Z] = E[X] + E[Y] = 3 + 2 = 5.$$

Thus the expected time that the professor leaves his office is 5 hours after 9 a.m.

c) We define the following additional random variables:

W = length of time between 9 a.m. and arrival of the Ph.D. student (uniformly distributed between 9 a.m. and 5 p.m.)
R = amount of time the student will spend with the professor, if he finds the professor (uniformly distributed between 0 and 1 hours).
T = amount of time the professor will spend with the student
Let also F be the event that the student finds the professor.

To find $E[T]$, we write

$$E[T] = P(F)E[T|F] + P(F^c)E[T|F^c]$$

Using the problem data,

$$E[T|F] = E[R] = \frac{1}{2}$$

(this is the expected value of a uniformly distribution ranging from 0 to 1),

$$E[T|F^c] = 0$$

(since the student leaves if he does not find the professor). We have

$$E[T] = E[T|F]P(F) = \frac{1}{2}P(F)$$

so we need to find $P(F)$.
In order for the student to find the professor, his arrival should be between the arrival and the departure of the professor. Thus

$$P(F) = P(Y \leq W \leq X + Y)$$

We have that W can be between 0 (9 a.m.) and 8 (5 p.m.), but $X + Y$ can be any value greater than 0. In particular, it may happen that the sum is greater than the upper bound for W. We write

$$P(F) = P(Y \leq W \leq X + Y) = 1 - (P(W < Y) + P(W > X + Y))$$

We have

$$P(W < Y) = \int_0^4 \frac{1}{4} \int_0^y \frac{1}{8} dwdy = \frac{1}{4}$$

and

$$\begin{aligned}
P(W > X+Y) &= \int_0^4 P(W > X+Y | Y = y) f_Y(y) dy \\
&= \int_0^4 P(X < W - Y | Y = y) f_Y(y) dy \\
&= \int_0^4 \int_y^8 F_{X|Y}(w-y) f_W(w) f_Y(y) dwdy \\
&= \int_0^4 \frac{1}{4} \int_y^8 \frac{1}{8} \int_0^{w-y} \frac{1}{5-y} e^{-\frac{x}{5-y}} dxdwdy \\
&= \frac{12}{32} + \frac{1}{32} \int_0^4 (5-y) e^{-\frac{8-y}{5-y}} dy
\end{aligned}$$

Integrating numerically, we have

$$\int_0^4 (5-y) e^{-\frac{8-y}{5-y}} dy = 1.7584$$

Thus,

$$P(Y \le W \le X+Y) = 1 - (P(W < Y) + P(W > X+Y)) = 1 - 0.68 = 0.32$$

The expected amount of time the professor will spend with the student is then

$$E[T] = \frac{1}{2} P(F) = \frac{1}{2} 0.32 = 0.16 = 9.6 \text{ mins.}$$

Next, we want to find the expected time the professor will leave his office. Let Z be the length of time measured from 9 a.m. until he leaves his office. If the professor doesn't spend any time with the student, then Z will be equal to $X + Y$. On the other hand, if the professor is interrupted by the student, then the length of time will be equal to $X + Y + R$. This is because the professor will spend the same amount of total time on the task regardless of whether he is interrupted by the student. Therefore,

$$E[Z] = P(F)E[Z|F] = P(F^c)E[Z|F^c] = P(F)E[X+Y+R] + P(F^c)E[X+Y]$$

Using the results of the earlier calculations,

$$E[X+Y] = 5$$

$$E[X+Y+R] = E[X+Y] + E[R] = 5 + \frac{1}{2} = \frac{11}{2}$$

Therefore,

$$E[Z] = 0.68 \cdot 5 + 0.32 \cdot \frac{11}{2} = 5.16$$

Thus the expected time the professor will leave his office is 5.16 hours after 9 a.m.

Problem 7 (Ch4 P35) Let X be a random variable that takes nonnegative integer values, and is associated with a transform of the form

$$M_x(s) = c.\frac{3 + 4e^{2s} + 2e^{3s}}{3 - e^s}$$

where c is some scalar. Find E[X], $P_x(1)$, and $E[X|X \neq 0]$.

Solution

We first find c by using the equation

$$1 = M_X(0) = c \cdot \frac{3 + 4 + 2}{3 - 1}$$

so that $c = 2/9$. We then obtain

$$E[X] = \frac{dM_X}{ds}(s)|_{s=0} = \frac{2}{9}\frac{(3 - e^s)(8e^{2s} + 6e^{3s}) + e^s(3 + 4e^{2s} + 2e^{3s})}{(3 - e^s)^2}|_{s=0} = \frac{37}{18}$$

We now use the identity

$$\frac{1}{3 - e^s} = \frac{1}{3}\frac{1}{1 - \frac{e^s}{3}} = \frac{1}{3}(1 + \frac{e^s}{3} + \frac{e^{2s}}{9} + ...).$$

By identifying the coefficients of e^{0s} and e^s we obtain

$$p_X(0) = \frac{2}{9}$$
$$p_X(1) = \frac{2}{27}$$

Let $A = \{X \neq 0\}$. We have

$$p_{X|\{X \in A\}} = \begin{cases} \frac{p_X(k)}{P(A)}, & \text{if } k \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

so that

$$\begin{aligned} E[X|X \neq 0] &= \sum_{k=1}^{\infty} k p_{X|A}(k) \\ &= \sum_{k=1}^{\infty} \frac{k p_X(k)}{P(A)} \\ &= \frac{E[X]}{1 - p_X(0)} \\ &= \frac{37/18}{7/9} \\ &= \frac{37}{14} \end{aligned}$$

Problem 8 (Ch4 P43) A motorist goes through 4 lights, each of which is found to be red with probability 1/2. The waiting times at each light are modeled as independent normal random variables with a mean 1 minute and a standard deviation 1 /2 minute. Let X be the total waiting time at the red lights.

(a) Use the total probability theorem to find the PDF and the transform associated with X, and the probability that X exceeds 4 minutes. Is X normal?

(b) Find the transform associated with X by viewing X as a sum of a random number of random variables.

Solution

a) Using the total probability theorem, we have

$$P(X > 4) = \sum_{k=0}^{4} P(k \text{ lights are red}) P(X > 4 | k \text{ lights are red})$$

We have

$$P(k \text{ lights are red}) = \binom{4}{k} \left(\frac{1}{2}\right)^4$$

The conditional PDF of X given that k lights are red, is normal with mean k minutes and standard deviation $(1/2)\sqrt{k}$. Thus, X is a mixture of normal random variables and the transform associated with its (unconditional) PDF is the corresponding mixture of the transforms associated with the (conditional) normal PDFs. However, X is not normal, because a mixture of normal PDFs need not be normal. The probability $P(X > 4|k \text{ lights are red})$ can be computed from the normal tables for each k, and $P(X > 4)$ is obtained by substituting the results in the total probability formula above.

b) Let K be the number of traffic lights that are found to be red. We can view X as the sum of K independent normal random variables. Thus the transform associated with X can be found by replacing in the binomial transform $M_K(s) = (1/2 + (1/2)e^s)^4$ the occurrence of e^s by the normal transform corresponding to $\mu = 1$ and $\sigma = 1/2$. Thus,

$$M_X(s) = \left(\frac{1}{2} + \frac{1}{2}\left(e^{\frac{(1/2)^2 s^2}{2} + s}\right)\right)^4$$

Note that by using the formula for the transform, we cannot easily obtain the probability $P(X > 4)$.

Problem 9 (Ch5 P5) Let $X_1, X_2, ...$ be independent random variables that are uniformly distributed over [-1,1]. Show that the sequence $Y_1, Y_2, ...$ Converges in probability to some limit, and identify the limit, for each of the following cases:

(a) $Y_n = X_n/n$

(b) $Y_n = (X_n)^n$

(c) $Y_n = X_1 X_2 ... X_n$

(d) $Y_n = max(X_1, ..., X_n)$

Solution

In cases (a), (b), and (c), we show that Y_n converges to 0 in probability. In case (d), we show that Y_n converges to 1 in probability.
(a) For any $\epsilon > 0$, we have

$$P(|Y_n| \geq \epsilon) = 0$$

for all n with $1/n < \epsilon$, so $P(|Y_n| \geq \epsilon) \to 0$.
(b) For all $\epsilon \in (0, 1)$, we have

$$P(|Y_n| \geq \epsilon) = P(|X_n|^n \geq \epsilon) = P(X_n \geq \epsilon^{1/n}) + P(X_n \leq -\epsilon^{1/n}) = 1 - \epsilon^{1/n}$$

(c) Since $X_1, X_2, ...$ are independent random variables, we have

$$E[Y_n] = E[X_1] \cdots E[X_n] = 0$$

Also

$$var(Y_n) = E[Y_n^2] = E[X_1^2] \cdots E[X_n^2] = var(X_1)^2 = (\frac{4}{12})^n$$

so $var(Y_n) \to 0$. Since all Y_n have 0 as a common mean, from Chebyshev inequality it follows that Y_n converges to 0 in probability.

(d) We have for $\epsilon \in (0, 1)$, using the independence of $X_1, X_2, ...$,

$$\begin{aligned} P(|Y_n - 1| \geq \epsilon) &= P(\max\{X_1, ..., X_n\} \geq 1 + \epsilon) + P(\max\{X_1, ..., X_n\} \leq 1 - \epsilon) \\ &= P(X_1 \leq 1 - \epsilon, ..., X_n \leq 1 - \epsilon) \\ &= (P(X_1 \leq 1 - \epsilon))^n \\ &= (1 - \frac{\epsilon}{2})^n \end{aligned}$$

Hence $P(|Y_n - 1| \geq \epsilon) \to 0$.

2 Extra Problems:

Problem 10 Besides the transform that you learned in Section 4.4, there are various other similar transforms (moment generating functions) that can be defined. For example, for a discrete random variable X whose possible values are nonnegative integers, we can define its Laplace moment generating function as

$$G_X(z) = \sum_{k=0}^{\infty} p(X = k) z^k \tag{1}$$

Clearly, it is the z-transform of the sequence of mass probabilities $p_k = P(X = k)$. Actually, if we set $z = e^s$, this definition goes back to the definition that you learned in Section 4.4. However, the Laplace moment generating function is easier to deal with than the general moment generating function when X can only take nonnegative integer values. In particular, it may help solve certain counting problems in a more "mechanical" manner by using formula of power series and therefore does not require much "cleverness" (see question (5) below). All series involved in this problem are assumed to converge.

(1) Show that $E(X) = \frac{dG_X(z)}{dz}|_{z=1}$ and $Var(X) = \frac{d^2G_X(z)}{dz^2}|_{z=1} + \frac{dG_X(z)}{dz}|_{z=1} - (\frac{dG_X(z)}{dz}|_{z=1})^2$

(2) For the Binomial random variable X (with parameters n and p), please compute its Laplace moment generating function and use results in (1) to compute its expected value and its variance.

(3) Show that if $X_1, \ldots, X_n$ are independent random variables and $X = \sum_{i=1}^{n} X_i$, then $G_X(z) = \prod_{i=1}^{n} G_{X_i}(z)$

(4) Please compute the Laplace moment generating function for the Bernoulli random variable (with parameter p). Then using the fact that the Binomial random variable X (with parameters n and p) can be expressed as the sum of n Bernoulli random variables (with parameter p) and the result in (3) to compute $G_X(z)$ and compare with your result in (2).

(5) You randomly roll five six-sided dice, what is the probability that the sum of those five numbers is 15? Hint: Doing counting directly could be quite involved here. Please compute the Laplace generating function of the random variable that is of your interest first. Also, you may find the following negative binomial series useful.

$$(1+z)^{-m} = \sum_{i=0}^{\infty} (-1)^i \binom{m+i-1}{i} z^i \tag{2}$$

Solution (1) We will first show that $E[X] = \frac{dG_X(z)}{dz}|_{z=1}$ by starting from the right hand side and ending with left hand side expression.

$$\begin{aligned}
\frac{dG_X(z)}{dz}|_{z=1} &= \sum_{k=0}^{\infty} P(X=k)\frac{dz^k}{dz}|_{z=1} \\
&= \sum_{k=0}^{\infty} P(X=k)kz^{k-1}|_{z=1} \\
&= \sum_{k=0}^{\infty} P(X=k)k \\
&= E[X]
\end{aligned}$$

To show $var(X) = \frac{d^2G_X(z)}{dz^2}|_{z=1} + \frac{dG_X(z)}{dz}|_{z=1} - (\frac{dG_X(z)}{dz}|_{z=1})^2$, again we will move from RHS to LSH, but first figure out what $\frac{d^2G_X(z)}{dz^2}|_{z=1}$ means in terms of moments of X.

$$\begin{aligned}
\frac{d^2G_X(z)}{dz^2}|_{z=1} &= \sum_{k=0}^{\infty} P(X=k)\frac{d^2z^k}{dz^2}|_{z=1} \\
&= \sum_{k=0}^{\infty} P(X=k)k(k-1)z^{k-2}|_{z=1} \\
&= \sum_{k=0}^{\infty} P(X=k)k(k-1) \\
&= E[X^2] - E[X]
\end{aligned}$$

Since we already found what $\frac{dG_X(z)}{dz}|_{z=1}$ is, the RHS becomes

$$\text{RHS} = (E[X^2] - E[X]) + E[X] - E[X]^2 = E[X^2] - E[X]^2 = var(X)$$

(2) First we compute the Laplace moment generating function $G_X(z)$.

$$\begin{aligned}
G_X(z) &= \sum_{k=0}^{\infty} P(X=k)z^k \\
&= \sum_{k=0}^{n} \binom{n}{k} p^k(1-p)^{n-k}z^k \\
&= (pz + (1-p))^n
\end{aligned}$$

where the last line comes from the fact that $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.

Now, first and second derivatives of $G_X(z)$ with respect to z are

$$\begin{aligned}
\frac{dG_X(z)}{dz}|_{z=1} &= np(pz+(1-p))^{n-1}|_{z=1} \\
&= n(p+(1-p))^{n-1} \\
&= np \\
\frac{d^2G_X(z)}{dz^2}|_{z=1} &= n(n-1)p^2(pz+(1-p))^{n-2}|_{z=1} \\
&= (np)^2 - np^2
\end{aligned}$$

Plugging these into expressions in (1) give

$$E[X] = np$$

$$var(X) = ((np)^2 - np^2) + np - (np)^2 = np(1-p)$$

(3) Approach I
To show that $G_X(z) = \prod_{i=1}^{n} G_{X_i}(z)$ given $X_1, ..., X_n$ are independent r.v.'s and $X = \sum_{i=1}^{n} X_i$, we first notice that $G_X(z) = \sum_{k=0}^{\infty} P(X=k)z^k = E[z^X]$. Then, using independence,

$$\begin{aligned}
G_X(z) &= E[z^X] \\
&= E[z^{\sum_{i=1}^{n} X_i}] \\
&= \prod_{i=1}^{n} E[z^{X_i}] \\
&= \prod_{i=1}^{n} G_{X_i}(z)
\end{aligned}$$

Approach II
Another approach to this problem is to use mathematical induction. Recall that mathematical induction has two parts to it - the "basis" and "inductive step".
i) Basis: $i = 1$

If $i = 1$, then $X = X_1$. Thus, the given expression becomes trivial

$$\begin{aligned}
LHS &= G_X(z) = G_{X_1}(z) \\
RHS &= \prod_{i=1}^{1} G_{X_i}(z) = G_{X_1}(z)
\end{aligned}$$

ii) Inductive Step: Assume that $G_{X'}(z) = \prod_{i=1}^{n-1} G_{X_i}(z)$ holds for $X' = \sum_{i=1}^{n-1} X_i$ and show that $G_X(z) = \prod_{i=1}^{n} G_{X_i}(z)$ is true for $X = \sum_{i=1}^{n} X_i$.

Since $X = X' + X_n$, $G_X(z)$ can be expressed as the following using convolution ($k' = k - m$).

$$\begin{aligned}
G_X(z) &= \sum_{k=0}^{\infty} P(X' + X_n = k) z^k \\
&= \sum_{k=0}^{\infty} \sum_{m=0}^{k} P(X' = m) P(X_n = k - m) z^k \\
&= \sum_{m=0}^{\infty} \sum_{k=m}^{\infty} P(X' = m) P(X_n = k - m) z^k \\
&= \sum_{m=0}^{\infty} P(X' = m) z^m \sum_{k=m}^{\infty} P(X_n = k - m) z^{k-m} \\
&= G_{X'}(z) \sum_{k'=0}^{\infty} P(X_n = k') z^{k'} \\
&= G_{X'}(z) G_{X_n}(z) \\
&= \prod_{i=1}^{n-1} G_{X_i}(z) \cdot G_{X_n}(z) \\
&= \prod_{i=1}^{n} G_{X_i}(z)
\end{aligned}$$

(4) To use the results of (3), we first have to find the $G_X(z)$ for Bernoulli r.v. X_i with parameter p which is

$$\begin{aligned}
G_{X_i}(z) &= \sum_{k=0}^{\infty} P(X = k) z^k \\
&= \sum_{k=0}^{1} P(X = k) z^k \\
&= (1 - p) \cdot z^0 + p \cdot z^1 \\
&= pz + (1 - p)
\end{aligned}$$

Now, using the result of (3) and the fact that sum of n independent Bernoulli r.v's X_1,

X_2, ... X_n with parameter p becomes Binomial r.v. X with parameters n and p,

$$
\begin{aligned}
G_X(z) &= \prod_{i=1}^{n} G_{X_i}(z) \\
&= \prod_{i=1}^{n} (pz + (1-p)) \\
&= (pz + (1-p))^n
\end{aligned}
$$

This agrees with what we found in (2).

(5) Define random variable Y_i, $i \in \{1, 2, 3, 4, 5\}$ to represent the outcome of each dice. We know that Y_i has PMF of

$$
p_Y(y) = \begin{cases} \frac{1}{6} & \text{if } y = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise} \end{cases}
$$

which gives its Laplace moment generating function $G_{Y_i}(z)$ as

$$
\begin{aligned}
G_{Y_i} &= \frac{1}{6}(z + z^2 + ... + z^6) \\
&= \frac{1}{6} \frac{z(1-z^6)}{1-z}
\end{aligned}
$$

The last line uses the Z-transform tricks you learned in ECE 2200 (or Taylor expansion if you never took ECE 2200). Now, define r.v. $Y = \sum_{i=1}^{5} Y_i$ as sum of the outcome of five rolls and use result in (3) to first find $G_Y(z)$, and then expand the terms using negative binomial series given in hint.

$$
\begin{aligned}
G_Y(z) &= \prod_{i=1}^{5} G_{Y_i} \\
&= \left(\frac{1}{6}\right)^5 \frac{z^5(1-z^6)^5}{(1-z)^5} \\
&= \left(\frac{1}{6}\right)^5 z^5(1-z^6)^5 \sum_{k=0}^{\infty} (-1)^k \binom{5+k-1}{k} z^k \\
&= \left(\frac{1}{6}\right)^5 z^5(1-z^6)^5 \sum_{k=0}^{\infty} (-1)^k \binom{4+k}{k} z^k \\
&= \left(\frac{1}{6}\right)^5 z^5 \sum_{i=0}^{5} (-1)^i \binom{5}{i} z^{6i} \sum_{k=0}^{\infty} (-1)^k \binom{4+k}{k} z^k
\end{aligned}
$$

Finally, since $P(Y = k)$ is represented by the coefficient of z^k term in the expansion of $G_Y(z)$, we need to find the coefficient of z^{15} in the above expression.
By carefully examining the terms in the summations, it can be seen that the only terms that matter inside the second summation are z^{10} (multiplied together with z^5 term in the first the summation) and z^4 (multiplied together with $z^5 \cdot (-\binom{5}{1} z^6) = -5z^{11}$ term in the first summation). Since z^{10} and z^4 in the second summation have coefficients of $\binom{14}{10}$ and $\binom{8}{4}$, the final z^{15} term has coefficient of 0.0837 as shown below.

$$\left(\frac{1}{6}\right)^5 \left(z^5 \binom{14}{10} z^{10} - 5z^{11} \binom{8}{4} z^4\right) = \left[\left(\frac{1}{6}\right)^5 \left(\binom{14}{10} - 5\binom{8}{4}\right)\right] z^{15} = 0.0837z^{15}$$

Thus, 0.0837 is the probability that sum of five numbers is 15.

Problem 11 (Portfolio Optimization) The Acahti Trust Company offers two investment instruments, the Lousy Money Market Account and the Risky Mutual Fund. After one year, the Lousy Money Market Account returns exactly your initial principal, without interest. After one year, the Risky Mutual Fund returns twice your initial principal with probability 0.5, and half your initial principal with probability 0.5. Suppose you start with one dollar and each year you rebalance your investments so that a fraction θ of your savings is in the mutual fund and a fraction $1 - \theta$ is in the money market account. Let X_n denote the mutual fund's return during year n, so that

$$P(X_n = 2) = P(X_n = 1/2) = 1/2.$$

Let W_n denote your savings after n years.

(a) Express W_n in terms of W_{n-1}, X_n, and θ for $n \geq 1$.

(b) Determine $E[W_n]$ as a function of θ. What value of θ maximizes your expected savings?

(c) Show that

$$\lim_{n\to\infty} \frac{1}{n} \log_e(W_n) = \beta$$

with probability 1 for some constant β, and determine β as a function of θ. We call β the *asymptotic growth rate* of the savings because asymptotically W_n behaves like $\exp(\beta n)$ with probability 1.

Hint: Write W_n as a product of i.i.d. variables, take logarithms, and use the strong law of large numbers.

(d) What value of θ maximizes the asymptotic growth rate? What is this maximum growth rate? What asymptotic growth rate is achieved with the θ you found in (b)?

(e) How come trying to maximize average return is different from maximizing the long term return? Which quantity, the expected savings or the asymptotic growth rate, should you aim to optimize? Explain.

(a) Fraction θ goes into mutual fund whose return is X_n.
Fraction $(1-\theta)$ goes into money market account whose return is 1.

$$W_n = ((1-\theta) + \theta X_n)W_{n-1}$$

(b) Using independence between X_n and W_{n-1}, and using $E[X_n] = 2\cdot(1/2)+(1/2)\cdot(1/2) = 5/4$,

$$\begin{aligned}
E[W_n] &= E[((1-\theta) + \theta X_n)W_{n-1}] \\
&= ((1-\theta) + \theta E[X_n])E[W_{n-1}] \\
&= ((1-\theta) + \theta \frac{5}{4})E[W_{n-1}] \\
&= (1 + \frac{\theta}{4})E[W_{n-1}]
\end{aligned}$$

By induction, $E[W_n] = (1+\frac{\theta}{4})^n$. The expected savings $E[W_n]$ is therefore maximized when $\theta = 1$.

(c) Using what was found in (a), $W_n = ((1-\theta) + \theta X_n)W_{n-1}$,

$$\begin{aligned}
W_n &= ((1-\theta) + \theta X_n)W_{n-1} \\
&= ((1-\theta) + \theta X_n)((1-\theta) + \theta X_{n-1})W_{n-2} \\
&= ((1-\theta) + \theta X_n)((1-\theta) + \theta X_{n-1})...((1-\theta) + \theta X_1) \\
&= \prod_{i=1}^{n}((1-\theta) + \theta X_i)
\end{aligned}$$

Now, plug-in this expression into the given $\lim_{n\to\infty} \frac{1}{n}\ln(W_n)$, and use the fact that
1. X_i's have identical, independent distribution and
2. sample mean converges to true expectation as $n \to \infty$ to get

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\ln(W_n) &= \lim_{n\to\infty} \frac{1}{n}\ln(\prod_{i=1}^{n}((1-\theta) + \theta X_i)) \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n}\ln((1-\theta) + \theta X_i) \\
&= E[\ln((1-\theta) + \theta X_i)] \\
&= P(X_i = 2)\ln((1-\theta) + 2\theta) + P(X_i = 1/2)\ln((1-\theta) + \frac{\theta}{2}) \\
&= \frac{1}{2}(\ln(1+\theta) + \ln(1-\frac{\theta}{2})) \\
&= \beta
\end{aligned}$$

(d) You can plot out β to see that it is a concave function of θ. To find θ that maximizes the growth rate, take the derivative and set it equal to zero to get

$$\begin{aligned}
\frac{d\beta(\theta)}{\theta} &= 0 \\
\frac{1}{1+\theta} - \frac{1}{2-\theta} &= 0 \\
(2-\theta) - (1+\theta) &= 0 \\
\theta &= \frac{1}{2}
\end{aligned}$$

To find the maximum growth rate, plug in $\theta = 1/2$

$$\beta(\frac{1}{2}) = \frac{1}{2}(\ln(\frac{3}{2}) + \ln(\frac{3}{4})) \simeq 0.0589$$

We saw in part (b) that $E[W_n]$ is maximized for $\theta = 1$. Plugging in this value gives

$$\beta(1) = \frac{1}{2}(\ln(1+1) + \ln(1-\frac{1}{2})) = 0$$

(e) Any sensible argument will receive full credit. Here is one. Consider when $n = 2$. With equal probability, you may have bad year (B) when your mutual fund cuts to half or a good year (G) when your fund doubles. For each possible case, the overall return W_n based on the choice of θ is as follows

	$\theta = 1$	$\theta = \frac{1}{2}$
BB	1/4	9/16
BG	1	9/8
GB	1	9/8
GG	4	9/4

which shows that the expected return is larger when you choose to optimize the expected savings. However, the result maybe misleading because this strategy is benefiting from a huge gain (4) that occurs with small probability (GG). On the other hand, the strategy of optimizing the growth rate may have less expected return, but has a more steady distribution of return whether you go through a bad or good year. This trend holds true for general n number of years.

When the mutual fund behaves as we expect, with half good years and half bad, optimizing the growth rate maximizes the return. It accomplishes this by constantly rebalancing the portfolio, moving money into the mutual fund after a bad year and drawing money out after a good year.

On the other hand, the strategy to optimize the expectation is “chasing” the huge returns to be had in the unlikely event that the mutual fund has an abnormally large number of good years. It accomplishes this by leaving all of the money in the mutual fund and hoping for miracle.

Problem 12 Suppose that there are N students in ECE 4110. Sadly, when Dr. Tang returns the graded exam, he simply hands each student a random exam from the pile. Let K denote the number of students who, by chance, happen to receive their own exam. Thus N-K students receive someone else's exam.

(a) Use Markov's inequality to upper bound the chance that $K \geq m$, where $m > 1$.

(b) Use Chebychev's inequality to upper bound the chance that $K \geq m$, where $m > 1$.

Solution

(a) To use the Markov Inequality, we need to first find $E[K]$.
Let us represent the event that i^{th} person receives an exam with X_i. Then each X_1, X_2,... X_N are (not independent but) identically distributed with PMF of

$$p_{X_i}(x) = \begin{cases} 1/N & \text{if } x = 1 \\ (N-1)/N & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

where $x = 1$ represents the case when i^{th} person receives his own exam ($x = 0$ represents when i^{th} person does not receive his exam). To see why the PMF is formed this way, think about the probability that the first person receives his exam. This is $P(X_1 = 1) = 1/N$. By symmetry, $P(X_1 = 1) = P(X_2 = 1) = ... = P(X_N = 1) = 1/N$.
Now, since $K = X_1 + X_2 + ... + X_N$ from our definition, we can find $E[K]$ as

$$\begin{aligned} E[K] &= E[X_1 + X_2 + ... + X_N] \\ &= E[X_1] + E[X_2] + ...E[X_N] \\ &= N \cdot 1/N \\ &= 1 \end{aligned}$$

With $E[K]$ known, we can use the Markov Inequality to find the upper bound as

$$P(K \geq m) \leq \frac{E[K]}{m} = \frac{1}{m}$$

(b) To use the Chebysehv Inequality, we need to first find $var(K)$. To do so, I will first find the variance of X_i and covariance between X_i and X_j for $i \neq j$.
• Finding $var(X_i)$...

$$\begin{aligned} var(X_i) &= E[X_i^2] - E[X_i]^2 \\ &= \frac{1}{N} - (\frac{1}{N})^2 \\ &= \frac{N-1}{N^2} \end{aligned}$$

• Finding $cov(X_i, X_j)$, $(i \neq j)$...
First notice that

$$\begin{aligned} P(X_i X_j = 1) &= P(X_i = 1 \cap X_j = 1) \\ &= P(X_i = 1) P(X_j = 1 | X_i = 1) \\ &= \frac{1}{N} \frac{1}{N-1} \end{aligned}$$

where $P(X_j = 1|X_i = 1)$ comes from the fact that if i^{th} person has received his own exam, j^{th} person's exam is contained inside the remaining stack of $N-1$ exams.

Now, since $X_i X_j$ can only take values 0 or 1, find $cov(X_i, X_j)$ as

$$\begin{aligned} cov(X_i, X_j) &= E[X_i X_j] - E[X_i]E[X_j] \\ &= \frac{1}{N(N-1)} - (\frac{1}{N})^2 \\ &= \frac{1}{N^2(N-1)} \end{aligned}$$

• Finding $cov(K)$...
Since $K = X_1 + X_2 + ... + X_N$,

$$\begin{aligned} var(K) &= var(\sum_{i=1}^{N} X_i) \\ &= \sum_{i=1}^{N} var(X_i) + \sum_{\{(i,j)|i \neq j\}} cov(X_i, X_j) \\ &= N \cdot \frac{N-1}{N^2} + N(N-1) \cdot \frac{1}{N^2(N-1)} \\ &= 1 \end{aligned}$$

Finally, since K only takes non-negative integers, for $3 \leq m \leq N$, $|K - E[K]| \geq m - 1$ only holds when $K - 1 \geq m - 1$. This allows us to use the Chebyshev Inequality to find the upper bound as

$$\begin{aligned} P(|K - E[K]| \geq m - 1) &\leq \frac{var(K)}{(m-1)^2} \\ P(K - 1 \geq m - 1) &\leq \frac{1}{(m-1)^2} \\ P(K \geq m) &\leq \frac{1}{(m-1)^2} \end{aligned}$$

This part is optional for those who want to explore more. If you want to be rigorous, you may as well consider the tricky case when $m = 2$. For this case,

$$P(|K-1| \geq 1) = P(K \geq 2) + P(K = 0) \leq 1$$

which gives

$$P(K \geq 2) \leq 1 - P(K = 0)$$

where

$$P(K = 0) = \sum_{j=0}^{N} \frac{(-1)^j}{j!}$$