Northwestern University
Department of Electrical and Computer Engineering

ELEC ENG 422 Winter 2020

Problem set 2 Solutions:

Problems:

1. a.) X_2 and X_{10} are I.I.D. Bernoulli random variables and so their joint p.m.f. is simply

$$p_{X_2,X_{10}}(x_2,x_{10}) = \begin{cases} p^2, & \text{if } x_2 = x_{10} = 1, \\ p(1-p), & \text{if } (x_2 = 1 \text{ and } x_{10} = 0) \text{ or } (x_2 = 0 \text{ and } x_{10} = 1), \\ (1-p)^2, & \text{if } x_2 = x_{10} = 0, \\ 0, & \text{otherwise.} \end{cases}$$

b.) Z_t will be a geometric random variable (just like the interarrival times discussed in class). Hence,

$$p_{Z_t}(z) = \begin{cases} (1-p)^{z-1}(p), & \text{if } z = 1, 2, 3, \ldots \\ 0, & \text{otherwise.} \end{cases}$$

c.) If $W_t = w$, then looking backward intime it means that there were $w-1$ time instants with no arrivals and then one with an arrival. Since the arrivals are independent, this is again a geoemetric random variable with the same p.m.f. as Z_t in the previous part.

d.) Note that W_t is a function of the arrivals before time t and Z_t is a function of th earrivals after time t, since these are independent sets of arrivals, then W_t and Z_t will be independent.

3. a.) Consider the case where X is a continuous zero mean random variable. Since it has a finite expectation, it must be that $\int_{-\infty}^{\infty} x f_X(x)\, dx$ converges. For the variance to be infinite, it must be that $\int_{-\infty}^{\infty} x^2 f_X(x)\, dx$ does not converge. A sufficient condition for satisfying both of these is that $f_X(x) \approx 1/|x|^3$ as x goes to ∞ and $-\infty$. This means that $x f_X(x)$ would decrease like $1/x^2$, which is integrable, but $x^2 f_X(x)$ would decrease like $1/x$ which is not.

b.) No - if a random variable has a finite second moment its first moment must also be finite. One way to argue this for continuous random variables is that the "tails" of $x f_X(x)$ must decrease faster than the tails fo $x^2 f_X(x)$ and so if the later is integrable, the former must be too. Another way to show this is to use the Cauchy-Schwartz inequality.

4. Using linearity of expectations we have:

$$E[(X-z)^2] = E(X^2) - 2zE(X) + z^2$$

The right-hand side of this is a convex quadratic function of z and so can be minimized by setting its derivative equal to zero. Doing this and solving for z gives the desired result.

Note: this has a natural physical analog: one can view $E(X-z)^2$ as analogous to the moment of inertia of a solid about z - this is minimized when z is the center of mass, which corresponds to $E(X)$.

6.

a.) Let A be the event: 'X_1 is even.' One way to solve this is condition on the possible values of X_1 and then use iterative expectation, i.e.,

$$E(X|A) = \sum_{i=2,4,6} E(X|A, X_1 = i)P(X_1 = i|A),$$

where $P(X_1 = i|A) = 1/3$.

To calculate $E(X|A, X_1 = i)$ note that

$$P(X = k|A \text{ and } X_1 = i) = \begin{cases} 1/6, & \text{for } k = 1+1, i+2, \ldots, i+6 \\ 0, & \text{otherwise.} \end{cases}$$

and so $E(X|A, X_1 = i) = i + 3.5$. Finally, $E(X|A) = (1/3)(5.5 + 7.5 + 9.5) = 7.5$.

b.) In this case note that if $X = 9$, then the only possible value of X_1 are $3, 4, 5, 6$. These are all equally likely conditioned on $X = 9$ and so $E(X_1|X = 9) = 4.5$.

8. a.) For any number m, let X be a random variable which takes on the two values km and 0, with probabilities $1/k$ and $1 - 1/k$, respectively. Note that $E(X) = (1/k)(km) + (1 - 1/k)(0) = m$ and thus this clearly satisfies $\Pr(X \geq kE(X)) = \Pr(X \geq km) = 1/k$.

b.) Note that

$$\begin{aligned} \Pr(X \geq kE(X)) &= \Pr(X - E(X) \geq (k-1)E(X)) \\ &= \Pr(|X - E(X)| \geq (k-1)E(X)) \\ &\leq \frac{Var(X)}{(k-1)^2(E(X))^2} \end{aligned}$$

where the second line follows since X is non-negative and the third line follows from Chebyshev's inequality. For the given random variable in part (a), $Var(X) = \frac{1}{k}((k-1)^2(E(X))^2) + (1 - \frac{1}{k})(E(X))^2$. Substituting this into the above bound and simplifying, we have

$$\Pr(X \geq kE(X)) \leq \frac{1}{k} + \frac{1}{k(k-1)}$$

which is clearly larger than the bound given by Markov's inequality.

c.) The key here is that the argument in class applied to a fixed random variable as we increased k. Here, as k changes in part (a), we are changing the random variable. To make this more precise, consider a random variable X as in part (a) for a particular integer $\tilde{k} > 2$. Then for $k = \tilde{k}$, Markov's inequality does give a tighter bound than Chebyshev's on the probability that $X \geq kE(X)$, but if we increase k and do not change the random variable, eventually Chebyshev's inequality will give a better bound.

9. First note that $|X| \geq 10$ if and only if $X^4 \geq 10^4$, and so

$$\Pr(|X| \geq 10) = \Pr(X^4 \geq 10^4)$$

Next, as in the proof of the Chebyshev bound in class, we can use that X^4 is non-negative, to apply the Markov inequality giving

$$\Pr(X^4 \geq 10^4) \leq \frac{E(X^4)}{10^4} = \frac{30}{10^4} = .03.$$

Exercise 1.6: Show that for a continuous nonnegative rv X,

$$\int_0^\infty \Pr\{X > x\}\, dx = \int_0^\infty x \mathsf{f}_X(x)\, dx. \tag{A.2}$$

Hint 1: First rewrite $\Pr\{X > x\}$ on the left side of (A.2) as $\int_x^\infty \mathsf{f}_X(y)\, dy$. Then think through, to your level of comfort, how and why the order of integration can be interchnged in the resulting expression.

Solution: We have $\Pr\{X > x\} = \int_x^\infty \mathsf{f}_X(y)\, dy$ from the definition of a continuous rv. We look at $\mathsf{E}[X] = \int_0^\infty \Pr\{X > x\}\, dx$ as $\lim_{a\to\infty} \int_0^a \mathsf{F}^\mathsf{c}(x)\, dx$ since the limiting operation $a \to \infty$ is where the interesting issue is.

$$\begin{aligned}
\int_0^a \mathsf{F}^\mathsf{c}(x)\, dx &= \int_0^a \int_x^\infty \mathsf{f}_X(y)\, dy\, dx \\
&= \int_0^a \int_x^a \mathsf{f}_X(y)\, dy\, dx + \int_0^a \int_a^\infty \mathsf{f}_X(y)\, dy\, dx \\
&= \int_0^a \int_0^y \mathsf{f}_X(y)\, dx\, dy + a\mathsf{F}_X^\mathsf{c}(a).
\end{aligned}$$

We first broke the integral on the right into two parts, one for $y < x$ and the other for $y \geq x$. Since the limits of integration on the first part were finite, they could be interchanged. The inner integral of the first part is $y\mathsf{f}_X(y)$, so

$$\lim_{a\to\infty} \int_0^a \mathsf{F}_X^\mathsf{c}(x)\, dx = \lim_{a\to\infty} \int_0^a y\mathsf{f}_X(y)\, dy + \lim_{a\to\infty} a\mathsf{F}_X^\mathsf{c}(a).$$

Assuming that $\mathsf{E}[X]$ exists, the integral on the left is nondecreasing in A and has the finite limit $\overline{X}$. The first integral on the right is also nondecreasing and upper bounded by the first integral, so it also has a limit. This means that $\lim_{a\to\infty} a\mathsf{F}_X^\mathsf{c}(a)$ must also have a limit, say β. Now if $\beta > 0$, then for any $\epsilon \in (0, a)$, $a\mathsf{F}_X(a) > \beta - \epsilon$ for all sufficiently large a. For all such a, then $\mathsf{F}_X^\mathsf{c}(a) > (\beta - \epsilon)/a$. This would imply that $\overline{X} = \int_0^\infty \mathsf{F}_X^\mathsf{c}(x)\, dx = \infty$, which is a contradiction. Thus $\beta = 0$, *i.e.*, $\lim_{a\to\infty} a\mathsf{F}_X^\mathsf{c}(a) = 0$, establishing (A.2) for the case where $\mathsf{E}[X]$ is finite. The case where $\mathsf{E}[X]$ is infinite is a minor perturbation.

The result that $\lim_{a\to\infty} aF^{\mathsf{c}}_X(a) = 0$ is also important and can be seen intuitively from Figure 1.3.

Hint 2: As an alternate approach, derive (A.2) using integration by parts.

Solution: Using integration by parts and being less careful,

$$\int_0^\infty d(x\mathsf{F}^{\mathsf{c}}_X(x)) = -\int_0^\infty x\mathsf{f}_X(x)\,dx + \int_0^\infty \mathsf{F}^{\mathsf{c}}_X(x)\,dx.$$

The left side is $\lim_{a\to\infty} a\mathsf{F}^{\mathsf{c}}_X(a) - 0\mathsf{F}_X(0)$ so this shows the same thing, again requiring the fact that $\lim_{a\to\infty} a\mathsf{F}^{\mathsf{c}}_X(a) = 0$ when $\mathsf{E}[X]$ exists.

Exercise 1.11: a) For any given rv Y, express $\mathsf{E}[|Y|]$ in terms of $\int_{y<0} \mathsf{F}_Y(y)\,dy$ and $\int_{y\geq 0} \mathsf{F}^{\mathsf{c}}_Y(y)\,dy$. Hint: Review the argument in Figure 1.4.

Solution: We have seen in (1.34) that

$$\mathsf{E}[Y] = -\int_{y<0} \mathsf{F}_Y(y)\,dy + \int_{y\geq 0} \mathsf{F}^{\mathsf{c}}_Y(y)\,dy.$$

Since all negative values of Y become positive in $|Y|$,

$$\mathsf{E}[|Y|] = +\int_{y<0} \mathsf{F}_Y(y)\,dy + \int_{y\geq 0} \mathsf{F}^{\mathsf{c}}_Y(y)\,dy.$$

To spell this out in greater detail, let $Y = Y^+ + Y^-$ where $Y^+ = \max\{0, Y\}$ and $Y^- = \min\{Y, 0\}$. Then $Y = Y^+ + Y^-$ and $|Y| = Y^+ - Y^- = Y^+ + |Y^-|$. Since $\mathsf{E}[Y^+] = \int_{y\geq 0} \mathsf{F}^{\mathsf{c}}_Y(y)\,dy$ and $\mathsf{E}[Y^-] = -\int_{y<0} \mathsf{F}_Y(y)\,dy$, the above results follow.

b) For some given rv X with $\mathsf{E}[|X|] < \infty$, let $Y = X - \alpha$. Using (a), show that

$$\mathsf{E}[|X-\alpha|] = \int_{-\infty}^{\alpha} \mathsf{F}_X(x)\,dx + \int_{\alpha}^{\infty} \mathsf{F}^{\mathsf{c}}_X(x)\,dx.$$

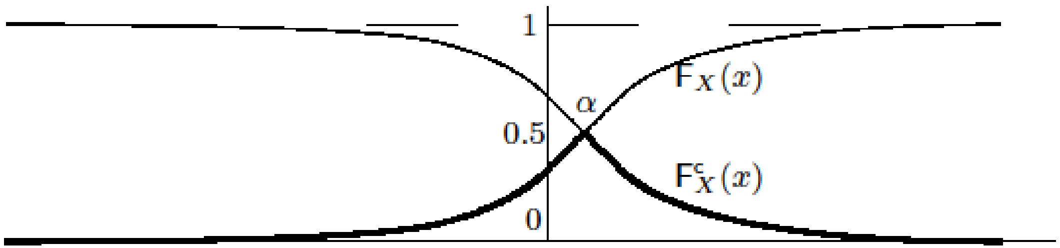
Solution: This follows by changing the variable of integration in (a). That is,

$$\begin{aligned}
\mathsf{E}[|X-\alpha|] &= \mathsf{E}[|Y|] = +\int_{y<0} \mathsf{F}_Y(y)\,dy + \int_{y\geq 0} \mathsf{F}^{\mathsf{c}}_Y(y)\,dy \\
&= \int_{-\infty}^{\alpha} \mathsf{F}_X(x)\,dx + \int_{\alpha}^{\infty} \mathsf{F}^{\mathsf{c}}_X(x)\,dx,
\end{aligned}$$

where in the last step, we have changed the variable of integration from y to $x - \alpha$.

c) Show that $\mathsf{E}[|X - \alpha|]$ is minimized over α by choosing α to be a median of X. Hint: Both the easy way and the most instructive way to do this is to use a graphical argument illustrating the above two integrals Be careful to show that when the median is an interval, all points in this interval achieve the minimum.

Solution: As illustrated in the picture, we are minimizing an integral for which the integrand changes from $\mathsf{F}_X(x)$ to $\mathsf{F}^{\mathsf{c}}_X(x)$ at $x = \alpha$. If $\mathsf{F}_X(x)$ is strictly increasing in x, then $\mathsf{F}^{\mathsf{c}}_X = 1 - \mathsf{F}_X$ is strictly decreasing. We then minimize the integrand over all x by choosing α to be the point where the curves cross, *i.e.*, where $\mathsf{F}_X(x) = .5$. Since the integrand has been minimized at each point, the integral must also be minimized.



If F_X is continuous but not strictly increasing, then there might be an interval over which $\mathsf{F}_X(x) = .5$; all points on this interval are medians and also minimize the integral; Exercise 1.10 (c) gives an example where $\mathsf{F}_X(x) = 0.5$ over the interval $[1, 2)$. Finally, if $\mathsf{F}_X(\alpha) \geq 0.5$ and $\mathsf{F}_X(\alpha - \epsilon) < 0.5$ for some α and all $\epsilon > 0$ (as in parts (a) and (b) of Exercise 1.10), then the integral is minimized at that α and that α is also the median.

Exercise 1.14: a) Let $X_1, X_2, \dots, X_n$ be rv's with expected values $\overline{X}_1, \dots, \overline{X}_n$. Show that $\mathsf{E}[X_1 + \cdots + X_n] = \overline{X}_1 + \cdots + \overline{X}_n$. You may assume that the rv's have a joint density function, but do not assume that the rv's are independent.

Solution: We assume that the rv's have a joint density, and we ignore all mathematical fine points here. Then

$$\begin{aligned}
\mathsf{E}[X_1+\cdots+X_n] &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(x_1+\cdots+x_n)\mathsf{f}_{X_1\cdots X_n}(x_1,\dots,x_n)\,dx_1\cdots dx_n \\
&= \sum_{j=1}^{n}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x_j\,\mathsf{f}_{X_1\cdots X_n}(x_1,\dots,x_n)\,dx_1\cdots dx_n \\
&= \sum_{j=1}^{n}\int_{-\infty}^{\infty} x_j\mathsf{f}_{X_j}(x_j)\,dx_j \;=\; \sum_{j=1}^{n}\mathsf{E}[X_j].
\end{aligned}$$

Note that the separation into a sum of integrals simply used the properties of integration and that no assumption of statistical independence was made.

b) Now assume that $X_1,\dots,X_n$ are statistically independent and show that the expected value of the product is equal to the product of the expected values.

Solution: From the independence, $\mathsf{f}_{X_1\cdots X_n}(x_1,\dots,x_n)=\prod_{j=1}^{n}\mathsf{f}_{X_j}(x_j)$. Thus

$$\begin{aligned}
\mathsf{E}[X_1X_2\cdots X_n] &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\prod_{j=1}^{n}x_j\prod_{j=1}^{n}\mathsf{f}_{X_j}(x_j)\,dx_1\cdots dx_n \\
&= \prod_{j=1}^{n}\int_{-\infty}^{\infty}x_j\,\mathsf{f}_{X_j}(x_j)\,dx_j \;=\; \prod_{j=1}^{n}\mathsf{E}[X_j].
\end{aligned}$$

c) Again assuming that $X_1,\dots,X_n$ are statistically independent, show that the variance of the sum is equal to the sum of the variances.

Solution: Since (a) shows that $\mathsf{E}\left[\sum_j X_j\right]=\sum_j\overline{X}_j$, we have

$$\begin{aligned}
\mathsf{VAR}\left[\sum_{j=1}^{n}X_j\right] &= \mathsf{E}\left[\left(\sum_{j=1}^{n}X_j-\sum_{j=1}^{n}\overline{X}_j\right)^2\right] \\
&= \mathsf{E}\left[\sum_{j=1}^{n}\sum_{i=1}^{n}(X_j-\overline{X}_j)(X_i-\overline{X}_i)\right] \\
&= \sum_{j=1}^{n}\sum_{i=1}^{n}\mathsf{E}\left[(X_j-\overline{X}_j)(X_i-\overline{X}_i)\right],
\end{aligned} \tag{A.5}$$

where we have again used (a). Now from (b) (which used the independence of the X_j), $\mathsf{E}\left[(X_j-\overline{X}_j)(X_i-\overline{X}_i)\right]=0$ for $i\neq j$. Thus(A.5) simplifies to

$$\mathsf{VAR}\left[\sum_{j=1}^{n}X_j\right] = \sum_{j=1}^{n}\mathsf{E}\left[(X_j-\overline{X}_j)^2\right] = \sum_{j=1}^{n}\mathsf{VAR}[X_j].$$