ECE 131A
Probability and Statistics
Instructor: Lara Dolecek

Homework 5 Solution
Monday, February 8, 2021
Due: Wednesday, February 17, 2021
before class begins

TAs: Lev Tauz levtauz@ucla.edu
Debarnab Mitra debarnabucla@g.ucla.edu

Please upload your homework to Gradescope by February 17, 3:59 pm.
Please submit a single PDF directly on Gradescope
You may type your homework or scan your handwritten version. Make sure all the work is discernible.

Chapters 4.6,4.7, and 4.9 of *Probability, Statistics, and Random Processes* by A. Leon-Garcia

1. Let X be the number of successes in n Bernoulli trials where the probability of success is p. Let $Y = \frac{X}{n}$ be the average number of successes per trial. Apply the Chebyshev inequality to the event $\{|Y - p| \geq a\}$. What happens as $n \to \infty$? What does this result imply about the distribution of Y as n goes to ∞?

   Solution:
   Let X be the number of successes in n Bernoulli trials and let $Y = X/n$. As such,

   $$E(Y) = \frac{E(X)}{n} = \frac{np}{n} = p$$

   $$Var(Y) = \frac{Var(X)}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

   Applying the Chebyshev inequality to the event $\{|Y - p| \geq a\}$, where $a > 0$, we have

   $$P\{|Y - p| \geq a\} \leq \frac{\sigma^2}{a^2} = \frac{p(1-p)}{na^2}.$$

   As $n \to \infty$, $P\{|Y - p| \geq a\} \to 0$ for any fixed $a > 0$.

   From this result, we can see that the probability that Y deviates from its mean p goes to zero as n goes to infinity. Thus, we can expect that Y converges to p.

2. Compare the Chebyshev inequality and the exact probability for the event $\{|X| > c\}$ as a function of c when X is a continuous uniform random variable in the interval $[-b, b]$ with $b > 0$. For $b = 1$, plot both the Chebyshev inequality and the exact probability for values of c that satisfy $0.2 \leq c \leq 1$.

   Solution:
   The Chebyshev inequality states that

   $$P\{|X - \mathbb{E}[X]| > c\} \leq \frac{\sigma^2}{c^2}.$$

Since X is a uniform random variable in the interval $[-b, b]$, we have

$$P(X \leq c) = \begin{cases} 0 & c \leq -b \\ \frac{c-b}{2b} & c \in [-b, b] \\ 1 & c \geq b \end{cases}$$

and $\mathbb{E}[X] = 0$.

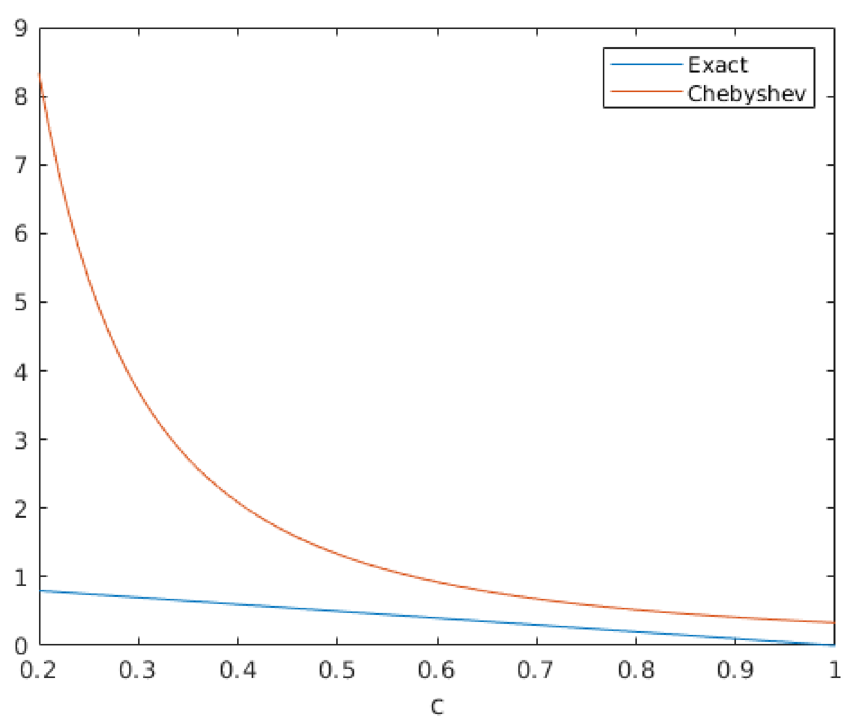
As such, for $0 \leq c \leq b$, we have

$$\begin{aligned} P\{|X| > c\} &= P(X > c) + P(X \leq -c) \\ &= [1 - P(X \leq c)] + P(X \leq -c) \\ &= [1 - \frac{c+b}{2b}] + \frac{-c+b}{2b} \\ &= 1 - \frac{c}{b} \end{aligned}$$

Now, we find the Chebyshev inequality. We again note that $E[X] = 0$ and $Var[X] = \frac{b^2}{3}$. Hence, for $0 \leq c \leq b$, we have

$$P\{|X| > c\} = P\{|X - \mathbb{E}[X]| > c\} \leq \frac{\sigma^2}{c^2} = \frac{b^2}{3c^2}$$

Now, we plot both the Chebyshev inequality and the exact probability.



We observe that the Chebyshev inequality largely overestimates the probability. For example, when $c \leq \frac{b}{\sqrt{3}}$ then the bound is always greater than 1 which is worthless for a probability value.

Additionally, we can use the Chebyshev inequality to prove some interesting relations between c and b:

$$\begin{aligned}
P\{|X| > c\} &\leq \frac{\sigma^2}{c^2} \\
1 - \frac{c}{b} &\leq \frac{b^2}{3} \times \frac{1}{c^2} \\
\frac{b-c}{b} &\leq \frac{b^2}{3c^2} \\
0 &\leq 3c^3 + b^3 - 3bc^2.
\end{aligned}$$

The inequality always holds for $0 \leq c \leq b$.

3. For each of the following random variables, find the characteristic function.

   (a) X is a discrete random variable that counts the number of failures when flipping a coin with success probability p before a success comes up. We can relate X to a geometric random variable A with parameter p by writing $X = A - 1$.

   Solution:

   The pmf for this random variable is

   $$P(X = k) = (1-p)^k p$$

   for $k \geq 0$ and is zero otherwise.

   As such,

   $$\begin{aligned}
   \Phi_X(w) = \mathbb{E}[e^{jwX}] &= \sum_{k=0}^{\infty} e^{jwk} P(X = k) \\
   &= p \sum_{k=0}^{\infty} e^{jwk} (1-p)^k \\
   &= p \sum_{k=0}^{\infty} (e^{jw}(1-p))^k \\
   &= \frac{p}{1 - (1-p)e^{jw}}.
   \end{aligned}$$

   (b) Y is a continuous uniform random variable and takes values in $[a, b]$ for real values a, b.

Solution:

$$\begin{aligned}\Phi_Y(w) = \mathbb{E}[e^{jwY}] &= \int_a^b e^{jwx} f(y)dy \\ &= \frac{1}{b-a}\int_a^b e^{jwy}dy \\ &= \frac{1}{b-a}\left(\frac{e^{jwy}}{jw}\Big|_a^b\right) \\ &= \frac{e^{jwb}-e^{jwa}}{jw(b-a)}\end{aligned}$$

(c) Z is a discrete uniform random variable where Y can only take values $\{c, c+1, c+2, \cdots, d\}$ for integers c, d.

Solution:

$$\begin{aligned}\Phi_Y(w) = \mathbb{E}[e^{jwZ}] &= \sum_{k=c}^{d} e^{jwk} P(Z=k) \\ &= \frac{1}{d-c+1}\sum_{k=c}^{d} e^{jwk} \\ &= \frac{1}{d-c+1}\cdot\frac{e^{jwc}-e^{jw(d+1)}}{1-e^{jw}}\end{aligned}$$

4. Let X be a continuous random variable. We define the random variable Y as $Y = aX+b$ for a, b such that $a > 0$. Let $\Phi_X(w)$ be the characteristic function of X.

(a) Let $\Phi_Y(w)$ be the characteristic function of Y. Determine $\Phi_Y(w)$ in terms of $\Phi_X(w)$, a, and b. Show the process for determining this relationship.

Solution:

$$\begin{aligned}\Phi_Y(w) = \mathbb{E}[e^{jwY}] &= \mathbb{E}[e^{jw(aX+b)}] \\ &= e^{jwb}\mathbb{E}[e^{jawX}] = e^{jwb}\Phi_X(aw)\end{aligned}$$

(b) Assume that $X \sim Exp(\lambda)$ and that $b = 0$. Using the characteristic function of Y, determine what kind of random variable Y is.

Solution:
From the previous part, we know that the characteristic function of $Y = aX$ is $\Phi_X(aw)$.
Recall from lecture that

$$\Phi_X(w) = \frac{\lambda}{\lambda - jw}$$

As such,

$$\Phi_Y(w) = \Phi_X(aw) = \frac{\lambda}{\lambda - jaw} = \frac{\lambda/a}{\lambda/a - jw}$$

By inspection, we see that Y is also a exponential random variable with parameter $\frac{\lambda}{a}$.

(c) Assume that $X \sim Exp(\lambda)$. First, use the properties of expectation to find $\mathbb{E}[Y]$. Now, use the characteristic function of Y to find $\mathbb{E}[Y]$. Answers are in terms of a, b, and λ.

Solution:
First, we use the properties of expectation. We note that $\mathbb{E}[X] = \frac{1}{\lambda}$. As such, we get

$$\mathbb{E}[Y] = \mathbb{E}[aX + b] = a \cdot \mathbb{E}[X] + b = a\frac{1}{\lambda} + b$$

Now, we use the characteristic function of Y to find $\mathbb{E}[Y]$. Recall that the moment theorem states that

$$\mathbb{E}[Y^n] = \frac{1}{j^n}\frac{d^n\Phi_Y(w)}{dw^n}|_{w=0}.$$

Again, recall from lecture that

$$\Phi_X(w) = \frac{\lambda}{\lambda - jw}$$

which results in

$$\Phi_Y(w) = \frac{e^{jwb}\lambda}{\lambda - jaw}.$$

Now we can use the moment theorem to find the expectation. First, we find the derivative of the characteristic function which is

$$\begin{aligned}\frac{d\Phi_Y(w)}{dw} &= \frac{d}{dw}(\frac{e^{jwb}\lambda}{\lambda - jaw}) \\ &= \lambda e^{jwb} \cdot \frac{jb(\lambda - jaw) + ja}{(\lambda - jaw)^2}.\end{aligned}$$

We can now calculate the expectation as follows

$$\begin{aligned}\mathbb{E}[Y] &= \frac{1}{j}\frac{d\Phi_Y(w)}{dw}|_{w=0} \\ &= \frac{\lambda}{j} \cdot \frac{jb\lambda + ja}{\lambda^2} \\ &= a\frac{1}{\lambda} + b\end{aligned}$$

which matches the result obtained by using the properties of expectation.

5. In this problem, you will be using the Chernoff bound to determine some useful upper bounds.

(a) Use the Chernoff bound to prove that $Q(x) \leq e^{\frac{-x^2}{2}} \quad \forall x \geq 0$. Recall that $Q(x) = P(X > x)$ where X is a normal distribution.

Solution:

Recall from discussion that the moment generating function of a Gaussian RV Z with mean μ and variance σ^2 is $M_Z(s) = e^{s\mu + \frac{\sigma^2 s^2}{2}}$.
Thus, we can apply the Chernoff bound on $Q(x)$ as follows:

$$\begin{aligned} Q(x) = P(X > x) &= P(X \geq x) && \text{X is continous} \\ &\leq \min_{s \geq 0} \frac{M_x(s)}{e^{sx}} && \text{Chernoff bound} \\ &= \min_{s \geq 0} \frac{e^{\frac{s^2}{2}}}{e^{sx}} \\ &= \min_{s \geq 0} e^{-sx + \frac{s^2}{2}} \end{aligned}$$

By taking the derivative, we get that $e^{-sx+\frac{s^2}{2}}(-x+s) = 0$ which means that the minimum is at $s = x$. Plugging this s back in gets us

$$Q(x) \leq e^{\frac{-x^2}{2}}$$

(b) Let $Y = \sum_{i=1}^{n} X_i$ where the X_i are i.i.d. Bernoulli random variables with parameter p. Using the Chernoff bound, prove that

$$P(Y \geq (1+\delta)np) \leq (\frac{e^{\delta}}{(1+\delta)^{1+\delta}})^{np} \quad \forall \delta > 0.$$

Hint: The following inequality may prove useful: $1 + y \leq e^y$ for all y.
What does this bound imply about the distribution of Y?

Solution:

The moment generating function of a Bernoulli random variable Z with parameter p is $M_Z(s) = e^s p + (1-p)$.
Recall that the moment generating function of a sum of independent RVs is equal to the product of all the individual moment generating functions of each RV. Thus,

$$
\begin{aligned}
P(Y \geq (1+\delta)np) &= \min_{s\geq 0} \frac{M_Y(s)}{e^{(1+\delta)nps}} \\
&= \min_{s\geq 0} \frac{\prod_{i=1}^{n} M_{X_i}(s)}{e^{(1+\delta)nps}} \\
&= \min_{s\geq 0} \frac{(e^s p + (1-p))^n}{e^{(1+\delta)nps}} \\
&\leq \min_{s\geq 0} \frac{(e^{p(e^s-1)})^n}{e^{(1+\delta)nps}} && \text{Using the hint inequality with } y = p(e^s-1) \\
&= \min_{s\geq 0} e^{np(e^s-1-s-s\delta)}
\end{aligned}
$$

By taking the derivative, we get the following equality $e^{np(e^s-1-s-s\delta)}(e^s - s + \delta) = 0$ which solving for s we get $s = \ln(1+\delta)$. Plugging this s back in, we get

$$
\begin{aligned}
P(Y \geq (1+\delta)np) &\leq e^{np(e^{\ln(1+\delta)}-1-\ln(1+\delta)-\ln(1+\delta)\delta)} \\
&= e^{np(\delta-\ln(1+\delta)(1+\delta))} \\
&= (\frac{e^\delta}{(1+\delta)^{1+\delta}})^{np}
\end{aligned}
$$

From this bound, we observe that the probability that Y is larger than $(1+\delta)\mathbb{E}[Y]$ gets smaller as the number of RVs increase. This indicates that the distribution concentrates more around the mean as n increases.

6. (Generating Random Variables)

In this question, you will use MATLAB to generate random variables using the transform method. Let U be a uniform random variable from 0 to 1. For $n > 0$, let $\{X_i \mid i \in [1, \cdots, n]\}$ be exponential random variables with parameter λ. Let $F(x)$ be the cdf of a single exponential random variable. You may use the MATLAB function *rand* to generate U.

(a) Determine the inverse of $F(x)$, i.e. $F^{-1}(u)$.

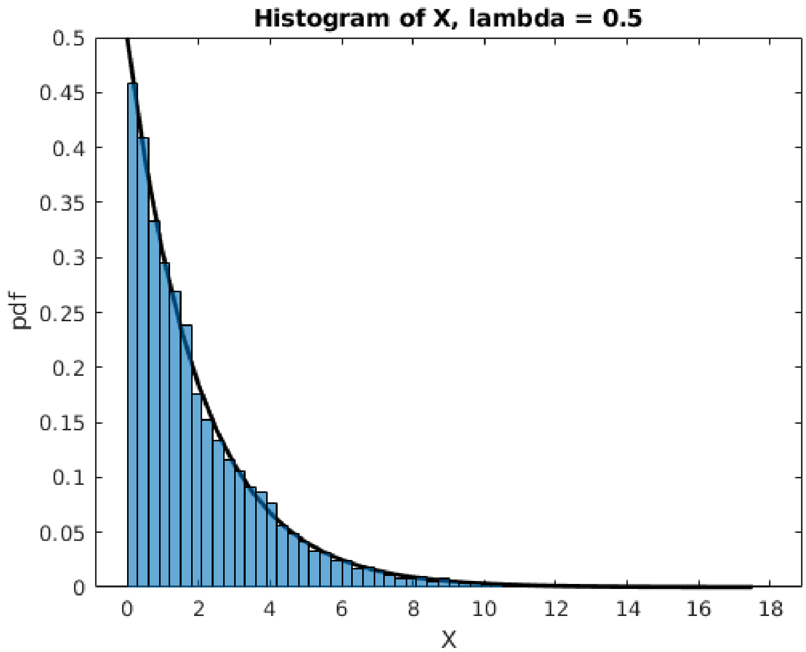
Solution:

$$F^{-1}(u) = -\frac{1}{\lambda}\ln(1-U)$$

(b) Now, you will generate X_i by applying $F^{-1}(u)$ on U. For $\lambda = 0.5$, generate $n = 5000$ exponential random variables and plot a histogram of the points. Compare the histogram to a plot of the pdf of X.

Solution:

As we would expect, the histogram matches very closely to the pdf of X.



(c) For $\lambda = 0.5$, generate $n = 5000$ exponential random variables and take the average of the random variables. How does it compare to the expectation of X?

Solution:

As you would expect, the average of X matches very closely to the expectation of X. I personally got 2.0168 which is very close to the expectation of 2. The reason this intuitive answer arises will be discussed later in the course when we go over law of large numbers.