CSE 312 Section 8

Tail Bounds, Joint Distributions, Law of Total Expectation

Announcements & Reminders

- HW6
  - Due on Wednesday 2/26
  - Late deadline Saturday 3/1 @ 11:59 pm

- HW7
  - Released
  - Due Wednesday 3/5 @ 11:59 pm
  - Late deadline Saturday 3/8 @ 11:59 pm

Review & Questions

Any lingering questions from this last week?

Each week in section, we’ll be reviewing the main concepts from this week and putting them into action by going through some practice problems together. But before we get into that review, we’ll try to start off each section with some time for you to ask questions. Was anything particularly confusing this week? Is there anything we can clarify before we dive into the review? This is your chance to clear things up!

Review of Main Concepts

- **Multivariate: Discrete to Continuous:**

	Discrete	**Continuous**
Joint PMF/PDF	$p_{X,Y}(x,y) = \mathbb{P}(X=x, Y=y)$	$f_{X,Y}(x,y) \neq \mathbb{P}(X=x, Y=y)$
Joint range/support $\Omega_{X,Y}$	$\{(x,y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x,y) > 0\}$	$\{(x,y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x,y) > 0\}$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \le x, s \le y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(t,s)\, ds dt$
Normalization	$\sum_{x,y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$
Expectation	$\mathbb{E}[g(X,Y)] = \sum_{x,y} g(x,y) p_{X,Y}(x,y)$	$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy$
Independence must have	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$

Review of Main Concepts

- **Law of Total Probability (r.v. version):** If X is a discrete random variable, then

$$\mathbb{P}(A) = \sum_{x \in \Omega_X} \mathbb{P}(A | X = x) p_X(x) \qquad \text{discrete } X$$

- **Law of Total Expectation (Event Version)**: Let X be a discrete random variable, and let events $A_1, ..., A_n$ partition the sample space. Then,

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X | A_i] \mathbb{P}(A_i)$$

- **Conditional Expectation**: See table. Note that linearity of expectation still applies to conditional expectation: $\mathbb{E}[X + Y | A] = \mathbb{E}[X | A] + \mathbb{E}[Y | A]$
- **Law of Total Expectation (RV Version)**: Suppose X and Y are random variables. Then,

$$\mathbb{E}[X] = \sum_{y} \mathbb{E}[X | Y = y] p_Y(y) \qquad \text{discrete version.}$$

Review of Main Concepts

- **Conditional distributions**

	Discrete	**Continuous**
Conditional PMF/PDF	$p_{X\|Y}(x\|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$	$f_{X\|Y}(x\|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
Conditional Expectation	$\mathbb{E}[X\|Y=y] = \sum_x x p_{X\|Y}(x\|y)$	$\mathbb{E}[X\|Y=y] = \int_{-\infty}^{\infty} x f_{X\|Y}(x\|y)dx$

- **Continuous Law of Total Probability:**

$$\mathbb{P}(A) = \int_{x \in \Omega_X} \mathbb{P}(A|X=x) f_X(x) dx$$

- **Continuous Law of Total Expectation:**

$$\mathbb{E}[X] = \int_{y \in \Omega_Y} \mathbb{E}[X|Y=y] f_Y(y) dy$$

Review of Main Concepts

- **Markov's Inequality**: Let X be a non-negative random variable, and $\alpha > 0$. Then,

$$\mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}[X]}{\alpha}$$

- **Chebyshev's Inequality**: Suppose Y is a random variable with $\mathbb{E}[Y] = \mu$ and $\mathrm{Var}(Y) = \sigma^2$. Then, for any $\alpha > 0$,

$$\mathbb{P}(|Y - \mu| \geq \alpha) \leq \frac{\sigma^2}{\alpha^2}$$

- **(Multiplicative) Chernoff Bound**: Let $X_1, X_2, ..., X_n$ be *independent* Bernoulli random variables.

  Let $X = \Sigma_{i=1}^{n} X_i$, and $\mu = \mathbb{E}[X]$. Then, for any $0 \leq \delta \leq 1$,

  - $\mathbb{P}(X \geq (1+\delta)\mu) \leq e^{-\frac{\delta^2 \mu}{3}}$
  - $\mathbb{P}(X \leq (1-\delta)\mu) \leq e^{-\frac{\delta^2 \mu}{2}}$

Review Questions

Question #1

a) True or False: Markov’s Inequality always gives a non-negative result.

- True
- False

Question #1

a) True or False: Markov's Inequality always gives a non-negative result.

- True
- False

Question #2

b) Suppose C and D are discrete random variables. Then $E[C|D=d] =$

- $\sum_d d \cdot p_{D|C}\,(d|c)$
- $\sum_c c \cdot p_{D|C}(d|c)$
- $\int_{-\infty}^{\infty} c f_{c|d} dx$
- $\frac{E[C]}{E[D]}$

Question #2

b) Suppose C and D are discrete random variables. Then $E[C|D=d] =$

- $\sum_d d \cdot p_{D|C}(d|c)$
- $\sum_c c \cdot p_{D|C}(d|c)$
- $\int_{-\infty}^{\infty} c f_{c|d} dx$
- $\frac{E[C]}{E[D]}$

Question #3

c) Suppose X and Y are random variables and A is an event. Given that $E[X|A] = 4$ and $E[Y|A] = 10$, what is $E[2X + \frac{Y}{2}|A]$

- 14
- 18
- 9
- 13

Question #3

c) Suppose X and Y are random variables and A is an event. Given that $E[X|A] = 4$ and $E[Y|A] = 10$, what is $E[2X + \frac{Y}{2}|A]$

- 14
- 18
- 9
- 13

Question #4

d) True or false: Chebyshev's Inequality can best be described as giving an upper bound on the distribution's right tail.

- True
- False

Question #4

d) True or false: Chebyshev's Inequality can best be described as giving an upper bound on the distribution's right tail.

- True
- False

Problem 6 –
Do You "Urn" to Learn More About Probability

Problem 6 – Do You “Urn” to Learn More About Probability

Suppose that 3 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let $X_i = 1$ if the i-th ball selected is white and let it be equal to 0 otherwise. Give the joint probability mass function of

a) X_1, X_2
b) X_1, X_2, X_3

Work on this with the people around you and then we’ll go over it together!

Problem 6 – Do You "Urn" to Learn More About Probability

a) X_1, X_2

Problem 6 – Do You "Urn" to Learn More About Probability

a) X_1, X_2

Here is one way of defining the joint pmf of X_1, X_2

$$\mathbb{P}(X_1 = 1, X_2 = 1) = \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 1 \mid X_1 = 1) = \frac{5}{13} \cdot \frac{4}{12} = \frac{20}{156}$$

$$\mathbb{P}(X_1 = 1, X_2 = 0) = \mathbb{P}(X_1 = 1)\mathbb{P}(X_2 = 0 \mid X_1 = 1) = \frac{5}{13} \cdot \frac{8}{12} = \frac{40}{156}$$

$$\mathbb{P}(X_1 = 0, X_2 = 1) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 1 \mid X_1 = 0) = \frac{8}{13} \cdot \frac{5}{12} = \frac{40}{156}$$

$$\mathbb{P}(X_1 = 0, X_2 = 0) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 0 \mid X_1 = 0) = \frac{8}{13} \cdot \frac{7}{12} = \frac{56}{156}$$

Problem 6 – Do You “Urn” to Learn More About Probability

b) X_1, X_2, X_3

Problem 6 – Do You “Urn” to Learn More About Probability

b) X_1, X_2, X_3

Instead of listing out all the individual probabilities, we could write a more compact formula for the pmf. In this problem, the denominator is always $P(13, k)$, where k is the number of random variables in the joint pmf. And the numerator is $P(5, i)$ times $P(8, j)$ where i and j are the number of 1s and 0s, respectively.

If we wish to compute $p_{X_1,X_2,X_3}(x_1, x_2, x_3)$, then the number of 1s (i.e., white balls) is $x_1 + x_2 + x_3$, and the number of 0s (i.e., red balls) is $(1 - x_1) + (1 - x_2) + (1 - x_3)$. Then, we can write the pmf as follows:

$$p_{X_1,X_2,X_3}(x_1, x_2, x_3) = \frac{10!}{13!} \cdot \frac{5!}{(5 - x_1 - x_2 - x_3)!} \cdot \frac{8!}{(5 + x_1 + x_2 + x_3)!}$$

Problem 3 – Exponential Tail Bounds

Problem 3 – Exponential Tail Bounds

Let $X \sim \text{Exp}(\lambda)$ and $k > \frac{1}{\lambda}$.

a) Use Markov's inequality to bound $P(X \geq k)$.
b) Use Markov's inequality to bound $P(X < k)$.
c) Use Chebyshev's inequality to bound $P(X \geq k)$.
d) What is the exact formula for $P(X \geq k)$.
e) For $\lambda k \geq 3$, how do the bounds given in parts a, c, and d compare?

Work on this with the people around you and then we'll go over it together!

Problem 3 – Exponential Tail Bounds

a) Use Markov's inequality to bound $P(X \geq k)$.

Problem 3 – Exponential Tail Bounds

a) Use Markov's inequality to bound $P(X \geq k)$.

We can use Markov's inequality here because X is non-negative since it is an exponential distribution. We also know that $E[X] = \lambda$ because $X \sim \text{Exp}(\lambda)$. By Markov's inequality, we get that:

$$\mathbb{P}(X \geq k) \leq \frac{1}{\lambda k}$$

Problem 3 – Exponential Tail Bounds

b) Use Markov's inequality to bound $P(X < k)$.

Problem 3 – Exponential Tail Bounds

b) Use Markov's inequality to bound $P(X < k)$.

From Markov's inequality (and our answer in (a)), we know that $P(X \geq k) \leq \frac{1}{\lambda k}$. Then,

$$\begin{aligned}
P(X \geq k) &\leq \frac{1}{\lambda k} \\
-P(X \geq k) &\geq -\frac{1}{\lambda k} && \text{multiplying be a negative flips the inequality} \\
1 - P(X \geq k) &\geq 1 - \frac{1}{\lambda k} \\
P(X < k) &\geq 1 - \frac{1}{\lambda k} && \text{by definition of complement}
\end{aligned}$$

Note that because we took the complement and the sign flipped, we have now found a *lower* bound for $P(X < k)$.

Problem 3 – Exponential Tail Bounds

c) Use Chebyshev's inequality to bound $P(X \geq k)$.

Problem 3 – Exponential Tail Bounds

c) Use Chebyshev's inequality to bound $P(X \geq k)$.

We rearrange algebraically to get into the form to apply Chebyshev's inequality. We then plug in the corresponding values and $Var(X) = \frac{1}{\lambda^2}$.

$$\mathbb{P}(X \geq k) = \mathbb{P}\left(X - \frac{1}{\lambda} \geq k - \frac{1}{\lambda}\right) \leq \mathbb{P}\left(\left|X - \frac{1}{\lambda}\right| \geq k - \frac{1}{\lambda}\right) \leq \frac{1}{\lambda^2(k - 1/\lambda)^2} = \frac{1}{(\lambda k - 1)^2}$$

Problem 3 – Exponential Tail Bounds

d) What is the exact formula for $P(X \geq k)$.

Problem 3 – Exponential Tail Bounds

d) What is the exact formula for $P(X \geq k)$.

Using the CDF for an exponential distribution and definition of complement:

$$\mathbb{P}(X \geq k) = 1 - P(X \leq k) = 1 - (1 - e^{-\lambda k}) = e^{-\lambda k}$$

Problem 3 – Exponential Tail Bounds

e) For $\lambda k \geq 3$, how do the bounds given in parts a, c, and d compare?

Problem 3 – Exponential Tail Bounds

e) For $\lambda k \geq 3$, how do the bounds given in parts a, c, and d compare?

$$e^{-\lambda k} < \frac{1}{(\lambda k - 1)^2} < \frac{1}{\lambda k}$$

so Markov's inequality gives the worst bound.

Problem 10 – Lemonade Stand

Problem 10 – Lemonade Stand

- Suppose I run a lemonade stand, which costs me $100 a day to operate. I sell a drink of lemonade for $20. Every person who walks by my stand either buys a drink or doesn't (no one buys more than one). If it is raining, n_1 people walk by my stand, and each buys a drink independently with probability p_1. If it isn't raining, n_2 people walk by my stand, and each buys a drink independently with probability p_2. It rains each day with probability p_3, independently of every other day. Let X be my profit over the next week. In terms of $n_1, n_2, p_1, p_2,$ and p_3, what is $E[X]$?

Work on this with the people around you and then we'll go over it together!

Problem 10 – Lemonade Stand

Let R be the event it rains. Let X_i be how many drinks I sell on day i for $i = 1, ..., 7$. We are interested in $X = \sum_{i=1}^{7} (20X_i - 100)$. We have $X_i|R \sim \text{Binomial}(n_1, p_1)$, so $\mathbb{E}[X_i|R] = n_1 p_1$. Similarly, $X_i|R^C \sim \text{Binomial}(n_2, p_2)$, so $\mathbb{E}\left[X_i|R^C\right] = n_2 p_2$. By the law of total expectation,

$$\mu = \mathbb{E}[X_i] = \mathbb{E}[X_i|R]\,\mathbb{P}(R) + \mathbb{E}\left[X_i|R^C\right]\mathbb{P}(R^C) = n_1 p_1 p_3 + n_2 p_2 (1 - p_3)$$

Problem 10 – Lemonade Stand

Let R be the event it rains. Let X_i be how many drinks I sell on day i for $i = 1, ..., 7$. We are interested in $X = \sum_{i=1}^{7} (20X_i - 100)$. We have $X_i | R \sim \text{Binomial}(n_1, p_1)$, so $\mathbb{E}[X_i|R] = n_1 p_1$. Similarly, $X_i|R^C \sim \text{Binomial}(n_2, p_2)$, so $\mathbb{E}\left[X_i|R^C\right] = n_2 p_2$. By the law of total expectation,

$$\mu = \mathbb{E}[X_i] = \mathbb{E}[X_i|R]\,\mathbb{P}(R) + \mathbb{E}\left[X_i|R^C\right]\mathbb{P}(R^C) = n_1 p_1 p_3 + n_2 p_2 (1 - p_3)$$

Hence, by linearity of expectation,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{7} (20X_i - 100)\right] = 20 \sum_{i=1}^{7} \mathbb{E}[X_i] - 700 = 140\mu - 700$$

$$= 140 \cdot (n_1 p_1 p_3 + n_2 p_2 (1 - p_3)) - 700.$$

Problem 9 – Trapped Miner

Problem 9 – Trapped Miner

A miner is trapped in a mine containing 3 doors.

- D_1: The 1st door leads to a tunnel that will take him to safety after 3 hours.
- D_2: The 2nd door leads to a tunnel that returns him to the mine after 5 hours.
- D_3: The 3rd door leads to a tunnel that returns him to the mine after a number of hours that is Binomial with parameters $(12, \frac{1}{3})$.

At all times, he is equally likely to choose any one of the doors. What is the expected number of hours for this miner to reach safety?

Problem 9 – Trapped Miner

Let T = number of hours for the miner to reach safety. (T is a random variable)
Let D_i be the event the i^{th} door is chosen. $i \in \{1, 2, 3\}$. Finally, let T_3 be the time it takes to return to the mine in the third case only (a random variable). Note that the expectation of T_3 is $12 * \frac{1}{3}$ because it is binomially distributed with parameters $n = 12, p = \frac{1}{3}$. By Law of Total Expectation, linearity of expectation, and by applying the conditional expectations given by the problem statement:

$$\begin{aligned}
\mathbb{E}[T] &= \mathbb{E}[T|D_1]\,\mathbb{P}(D_1) + \mathbb{E}[T|D_2]\,\mathbb{P}(D_2) + \mathbb{E}[T|D_3]\,\mathbb{P}(D_3) \\
&= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (\mathbb{E}[T_3 + T]) \cdot \frac{1}{3} \\
&= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (\mathbb{E}[T_3] + \mathbb{E}[T]) \cdot \frac{1}{3} \\
&= 3 \cdot \frac{1}{3} + (5 + \mathbb{E}[T]) \cdot \frac{1}{3} + (4 + \mathbb{E}[T]) \cdot \frac{1}{3}
\end{aligned}$$

Solving this equation for $\mathbb{E}[T]$, we get

$$\mathbb{E}[T] = 12$$

Therefore, the expected number of hours for this miner to reach safety is 12.

Problem 8 – Continuous Joint Density

Problem 8 – Continuous Joint Density

The joint density of X and Y is given by

$$f_{X,Y}(x,y) = \begin{cases} xe^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

And the joint density of W and V is given by

$$f_{W,V}(w,v) = \begin{cases} 2 & 0 < w < v, 0 < v < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent? Are W and V independent?

Problem 8 – Continuous Joint Density

For two random variables X, Y to be independent, we must have $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ for all $x \in \Omega_X,\ y \in \Omega_Y$. Let's start with X and Y by finding their marginal PDFs. By definition, and using the fact that the joint PDF is 0 outside of $y > 0$, we get:

$$f_X(x) = \int_0^{\infty} xe^{-(x+y)}dy = e^{-x}x$$

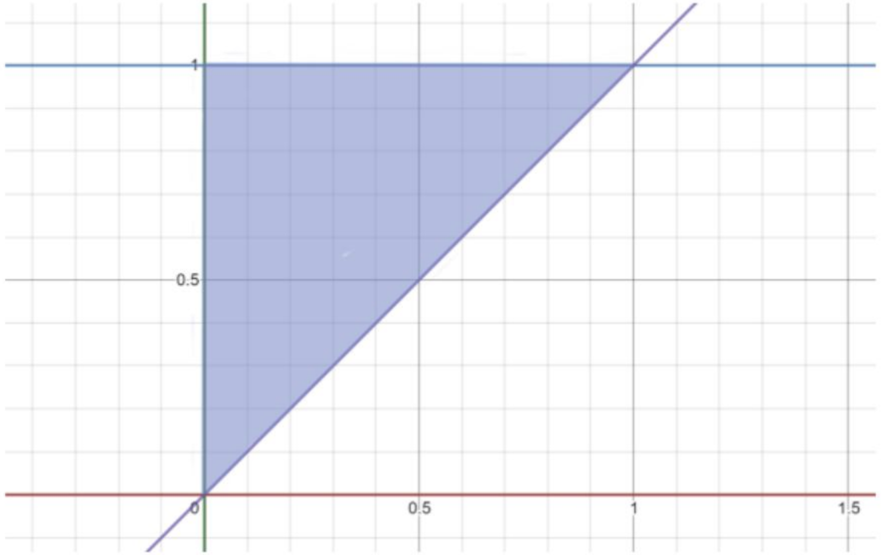
We do the same to get the PDF of Y, again over the range $x > 0$:

$$f_Y(y) = \int_0^{\infty} xe^{-(x+y)}dx = e^{-y}$$

Since $e^{-x}x \cdot e^{-y} = xe^{-x-y} = xe^{-(x+y)}$ for all $x, y > 0$, X and Y are independent.

Problem 8 – Continuous Joint Density

We can see that W and V are not independent simply by observing that $\Omega_W = (0,1)$ and $\Omega_V = (0,1)$, but $\Omega_{W,V}$ is not equal to their Cartesian product. Specifically, looking at their range of $f_{W,V}(w,v)$. Graphing it with w as the "x-axis" and v as the "y-axis", we see that :



The shaded area is where the joint pdf is strictly positive. Looking at it, we can see that it is not rectangular, and therefore it is not the case that $\Omega_{W,V} = \Omega_W \times \Omega_V$. Remember, the joint range being the Cartesian product of the marginal ranges is not sufficient for independence, but it is *necessary*. Therefore, this is enough to show that they are not independent.

That’s All, Folks!

Thanks for coming to section this week!
Any questions?