

Section 8: Tail Bounds, Joint Distributions, Law of Total Expectation (and bit of conditional distributions)

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	$F_{X,Y}(x,y) = \sum_{t \leq x, s \leq y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(t,s) ds dt$
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	$\forall x,y, p_{X,Y}(x,y) = p_X(x)p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$	$\forall x,y, f_{X,Y}(x,y) = f_X(x)f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$

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

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

- **Law of Total Expectation (Event Version):** Let X be a discrete random variable, and let events A_1, \dots, 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,

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

- **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:**

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

- **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 $\text{Var}(Y) = \sigma^2$. 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, \dots, X_n be *independent* Bernoulli random variables.

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

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

1. Content Review

- (a) True or false: the Union Bound always gives a result in $[0, 1]$.
- (b) True or false: Markov's Inequality always gives a non-negative result.
- (c) Suppose C and D are discrete random variables. Then $\mathbb{E}[C|D = d] =$
- $\sum_d dp_{D|C}(d|c)$
 - $\sum_c cp_{C|D}(c|d)$
 - $\int_{-\infty}^{\infty} cf_{c|d}dx$
 - $\frac{\mathbb{E}[C]}{\mathbb{E}[D]}$
- (d) Suppose X and Y are random variables and A is an event. Given that $\mathbb{E}[X|A] = 4$ and $\mathbb{E}[Y|A] = 10$, what is $\mathbb{E}[2X + Y/2|A]$?
- 14
 - 18
 - 9
 - 13
- (e) True or false: Chebyshev's Inequality can best be described as giving an upper bound on the distribution's right tail.

2. Tail bounds

Suppose $X \sim \text{Binomial}(6, 0.4)$. We will bound $\mathbb{P}(X \geq 4)$ using the tail bounds we've learned, and compare this to the true result.

- (a) Give an upper bound for this probability using Markov's inequality. Why can we use Markov's inequality?
- (b) Give an upper bound for this probability using Chebyshev's inequality. You may have to rearrange algebraically and it may result in a weaker bound.
- (c) Give an upper bound for this probability using the Chernoff bound.

(d) Give the exact probability.

3. Exponential Tail Bounds

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

(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?

4. Robbie's Late!

Suppose the probability Robbie is late to teaching lecture on a given day is at most 0.01. Do not make any independence assumptions.

(a) Use a Union Bound to bound the probability that Robbie is late at least once over a 30-lecture quarter.

(b) Use a Union Bound to bound the probability that Robbie is **never** late over a 30-lecture quarter.

(c) Use a Union Bound to bound the probability that Robbie is late at least once over a 120-lecture quarter.

5. Trinomial Distribution

A generalization of the Binomial model is when there is a sequence of n independent trials, but with three outcomes, where $\mathbb{P}(\text{outcome } i) = p_i$ for $i = 1, 2, 3$ and of course $p_1 + p_2 + p_3 = 1$. Let X_i be the number of times outcome i occurred for $i = 1, 2, 3$, where $X_1 + X_2 + X_3 = n$. Find the joint PMF $p_{X_1, X_2, X_3}(x_1, x_2, x_3)$ and specify its value for all $x_1, x_2, x_3 \in \mathbb{R}$.

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

7. Successes

Consider a sequence of independent Bernoulli trials, each of which is a success with probability p . Let X_1 be the number of failures preceding the first success, and let X_2 be the number of failures between the first 2 successes. Find the joint pmf of X_1 and X_2 . Write an expression for $E[\sqrt{X_1 X_2}]$. You can leave your answer in the form of a sum.

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?

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?

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 $\mathbb{E}[X]$?

11. 3 points on a line

Three points X_1, X_2, X_3 are selected at random on a line L (continuous independent uniform distributions). What is the probability that X_2 lies between X_1 and X_3 ?