# Section 7

## Review

- **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 | $\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$ | $\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$ |
| **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** | $\mathbb{E}[X\|Y=y] = \sum_x x \cdot p_{X\|Y}(x\|y)$ | $\mathbb{E}[X\|Y=y] = \int_{-\infty}^{\infty} x f_{X\|Y}(x\|y) dx$ |

- **Law of Total Probability (Continuous):** $A$ is an event, and $X$ is a continuous random variable with density function $f_X(x)$.

$$\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A \mid X = x) f_X(x) dx$$

- **Conditional Expectation**: Let $X$ and $Y$ be random variables. Then, the conditional expectation of $X$ given $Y = y$ is

$$\mathbb{E}[X|Y=y] = \sum_{x \in \Omega_X} x \cdot \mathbb{P}(X = x | Y = y) \qquad X \text{ discrete}$$

and for any event $A$,

$$\mathbb{E}[X|A] = \sum_{x \in \Omega_X} x \cdot \mathbb{P}(X = x | A) \qquad X \text{ discrete}$$

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

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

- **Law of Total Expectation (RV Version)**: Suppose $X$ and $Y$ are random variables. Then,

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

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} \mathbb{E}[X|Y=y] f_Y(y) dy \qquad Y \text{ continuous r.v.}$$

- **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 $\alpha > 0$,

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

- **(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 $0 \leqslant \delta \leqslant 1$,

  - $$\mathbb{P}\left(\sum_{i=1}^n X_i \geqslant (1+\delta)\mu\right) \leqslant \exp\left(-\frac{\delta^2 \mu}{3}\right)$$

  - $$\mathbb{P}\left(\sum_{i=1}^n X_i \leqslant (1-\delta)\mu\right) \leqslant \exp\left(-\frac{\delta^2 \mu}{2}\right)$$

## Task 1 – Content Review

**a)** Select one: For an event $A$ and a continuous random variable $X$ with density $f_X(x)$,

- ◯ $\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A \mid X = x)\mathbb{P}(X = x)dx$
- ◯ $\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A \mid X = x)f_X(x)dx$
- ◯ $\mathbb{P}(A) = \int_{-\infty}^{\infty} xf_X(x)dx$
- ◯ $\mathbb{P}(A) = \int_{-\infty}^{\infty} \mathbb{P}(A \mid X = x)dx$

**b)** True or false: the Union Bound always gives a result in $[0, 1]$.

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

**d)** Suppose $C$ and $D$ are discrete random variables. Then $\mathbb{E}[C|D = d] =$

- ◯ $\sum_d dp_{D|C}(d|c)$
- ◯ $\sum_c cp_{C|D}(c|d)$
- ◯ $\int_{-\infty}^{\infty} cf_{c|d}dx$
- ◯ $\frac{\mathbb{E}[C]}{\mathbb{E}[D]}$

**e)** 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

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

# Joint Distributions

## Task 2 – Who fails first?

Here's a question that commonly comes up in industry, but isn't immediately obvious. You have a disk with probability $p_1$ of failing each day. You have a CPU which independently has probability $p_2$ of failing each day. What is the probability that your disk fails *before* your CPU?

**a)** Compute the probability by summing over the relevant part of the probability space.

**b)** Try to provide an intuitive reason for the answer.

**c)** Recompute the probability using the law of total probability, conditioning on the value of $X_1$.

## Task 3 – 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?

# Conditional Distributions, Law of total expectation, Continuous LoTP

## Task 4 – A Dysfunctional Family

Rick and his grandson Morty are set to meet at a certain time. Since their relationship is a little strained, neither of them wants to be there on time. Let $X \sim Unif(0, 10)$ be the amount of minutes Morty is going to be late. Rick has cameras around the meeting spot and will observe Morty's arrival time $X = x$. Then, he will arrive at the meeting spot $Unif(x, 5x)$ minutes late. Let $Y$ be the random variable indicating how late Rick will be.

**a)** Using the above definitions determine $f_X$, $f_{Y|X}$, and $f_{XY}$. (You will want to determine $f_{YX}$ and use it to determine $f_{XY}$.).

**b)** Compute $\mathbb{E}[Y]$.

## Task 5 – Law of Total Probability Review

**a)** (Discrete version) Suppose we flip a coin with probability $U$ of heads, where $U$ is equally likely to be one of $\Omega_U = \{0, \frac{1}{n}, \frac{2}{n}, ..., 1\}$ (notice this set has size $n + 1$). Let $H$ be the event that the coin comes up heads. What is $\mathbb{P}(H)$?

**b)** Now suppose $U \sim$ Uniform(0,1) has the *continuous* uniform distribution over the interval $[0, 1]$. What is $\mathbb{P}(H)$?

Use the continuous version of the law of total probability: suppose $E$ is an event, and $X$ is a continuous random variable with density function $f_X(x)$. Then

$$\mathbb{P}(E) = \int_{-\infty}^{\infty} \mathbb{P}(E \mid X = x) f_X(x) dx$$

## Task 6 – 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 $X_3$?

## Task 7 – 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 $\mathbb{E}[X]$?

## Task 8 – Trapped Miner

A miner is trapped in a mine containing 3 doors.

- $D_1$: The $1^{st}$ door leads to a tunnel that will take him to safety after 3 hours.
- $D_2$: The $2^{nd}$ door leads to a tunnel that returns him to the mine after 5 hours.
- $D_3$: The $3^{rd}$ 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?

## Task 9 – Elevator rides

[This is the problem we did in class.] The number $X$ of people who enter an elevator on the ground floor is a Poisson random variable with mean 10. If there are N floors above the ground floor, and if each person is equally likely to get off at any one of the $N$ floors, independently of where others get off, compute the expected number of stops the elevator will make before discharging all the passengers.

# Tail Bounds

## Task 10 – Tail bounds

Suppose $X \sim \text{Binomial}(6, 0.4)$. We will bound $\mathbb{P}(X \geqslant 4)$ 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 a lower bound for $P(X < 4)$ using Markov's inequality.

**c)** 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.

**d)** Give an upper bound for this probability using the Chernoff bound.

**e)** Give an upper bound for $P(X \leqslant 2)$ using the Chernoff bound.

**f)** Give the exact probability of $P(X \geqslant 4)$.

## Task 11 – How many samples?

Let $X = X_1 + \ldots X_n$ be the sum of $n$ independent $Poisson(\lambda)$ random variables. Recall that the Poisson distribution has expectation and variance both equal to $\lambda$ and has the summation property that $X$ is a $Poisson(n\lambda)$ random variable.

**a)** How large a value of $n$ would Chebyshev's inequality need to guarantee that $\mathbb{P}(X \leqslant \mathbb{E}[X]/2) \leqslant 0.01$?

**b)** How large a value of $n$ would Markov's inequality need to guarantee that $\mathbb{P}(X \geqslant \mathbb{E}[X]/2) \leqslant 0.01$?

## Task 12 – Claris's Late!

Suppose the probability Claris 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 Claris is late at least once over a 30-lecture quarter.

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

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

## Task 13 – Exponential Tail Bounds

Let $X \sim \text{Exp}(\lambda)$ and $k > 1/\lambda$. Recall that $\mathbb{E}[X] = \frac{1}{\lambda}$ and $\text{Var}(X) = \frac{1}{\lambda^2}$.

**a)** Use Markov's inequality to bound $\text{P}(X \geqslant k)$.

**b)** Use Markov's inequality to bound $\text{P}(X < k)$.

**c)** Use Chebyshev's inequality to bound $\text{P}(X \geqslant k)$.

**d)** What is the exact formula for $\text{P}(X \geqslant k)$?

**e)** For $\lambda k \geqslant 3$, how do the bounds given in parts (a), (b), and (c) compare?