CSE 312 : Spring 2023 Final Exam

Name:	**NetID:** @uw.edu

Instructions

- You have 110 minutes to complete this exam.
- You are permitted one piece of 8.5x11 inch paper with handwritten notes (notes are allowed on both sides of the paper).
- You may not use a calculator or any other electronic devices during the exam.
- We will be scanning your exams before grading them. Please write legibly, and avoid writing up to the edge of the paper.
- If you run out of room, you may also use the last page for extra space, but tell us where to find your answer if it's not right below the problem.
- Since you don't have a calculator, you are generally free to **not** simplify expressions (though you may if you think it will be helpful).
- In general, show us the work you used to get to an answer, and explanations will help us reward partial credit, but we do not expect explanations at the level we usually require on homeworks.

Advice

- Writing a few words about where an expression came from is often very helpful for awarding partial credit.
- Remember to take deep breaths.

Question	**Max points**
Multiple Choice	30
Small Problems	11
Zoo	25
Cabbages	20
Concentration (Berries)	15
MLE (Baseball)	13
Morale	1
Total	**115**

This page was intentionally left blank.

1. [Multiple Choice] $0.25^{10} \approx 0.000001$ [30 points]

Each of the questions below has exactly one **correct** answer.
Fully fill in the circle of the best option below.

(a) A maximum likelihood estimator is

- ◯ Always consistent, but only sometimes unbiased.
- ◯ Always unbiased, but only sometimes consistent.
- ◯ Always both unbiased and consistent.
- ◯ None of the above.

(b) Let f_X be the pdf of a continuous random variable. Which of the following must be true?

- ◯ $\int_0^\infty f_X(z)dz = 1$
- ◯ $f_X(z) \leq 1$ for all z
- ◯ f_X is continuous
- ◯ None of the above.

(c) Which of the following is true of continuous random variables?

- ◯ $f_X(z)$ gives the probability that X takes on the value z.
- ◯ $\int_a^b f_X(z)\, dz$ gives $\mathbb{P}(a < X \leq b)$, for values $a < b$
- ◯ Conditioning is impossible for continuous random variables, because all outcomes have probability 0.

(d) How do you calculate the marginal value $p_X(x)$ from the joint pmf $p_{X,Y}(x, y)$?

- ◯ $\sum_x p_{X,Y}(x, y)$
- ◯ $\sum_y p_{X,Y}(x, y)$
- ◯ $\int_{-\infty}^{\infty} p_{X,Y}(x, y)dy$
- ◯ You cannot calculate the marginal from the joint pmf.

(e) We saw a "Las Vegas" algorithm; a version of quicksort that uses randomness. What is true about this algorithm?

- ◯ One should never use this algorithm as success cannot be guaranteed.
- ◯ One should run the algorithm (with independent randomness), repeatedly to increase success probability.
- ◯ One should run the algorithm (just once, as is).

(f) Cov$(X, Y) > 0$ indicates

- ◯ X and Y are independent.
- ◯ X being positive tends to indicate that Y is likely to be negative.
- ◯ X being more than its expectation tends to indicate that Y is likely to be more than its expectation.
- ◯ X being more than its expectation tends to indicate that Y is likely to be less than its expectation.

Each of the questions below has exactly one **correct** answer.
Fully fill in the circle of the best option below.

(g) Recall that $\Phi(z) = \mathbb{P}(Z \leq z)$, where $Z \sim \mathcal{N}(0, 1)$. If $k > 0$, select which of the following is equal to $\Phi(-k)$?

◯ $\Phi(k)$
◯ $1 - \Phi(k)$
◯ $\Phi(k + 1)$

(h) Suppose we have random samples $X_1, X_2, \ldots, X_n$ from a $\mathcal{N}(\theta, \sigma^2)$ random variable. Which of the following estimators is **biased**?

I. $\hat{\theta_1} = X_1$

II. $\hat{\theta_2} = \frac{1}{n-1} \sum_{i=1}^{n} X_i$

◯ $\hat{\theta_1}$ only
◯ $\hat{\theta_2}$ only
◯ $\hat{\theta_1}$ and $\hat{\theta_2}$
◯ Neither

(i) Let $X_1, X_2, \ldots, X_n$ be independent Bernoulli random variables, with $X = \sum_{i=1}^{n} X_i$ being the sum of these random variables. And suppose we know $\mathbb{E}[X]$ but do not know the parameters for each individual Bernoulli random variable. If we want to bound $P(X \geq t)$ for some fixed t, which of the following tail bounds can we **not** use here?

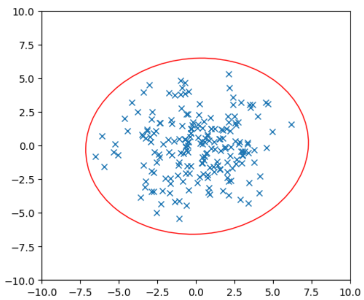
◯ Markov's inequality
◯ Chebyshev's inequality
◯ Chernoff bound
◯ More than one of the above are not usable here
◯ None of the above; all of the above tail bounds are usable here

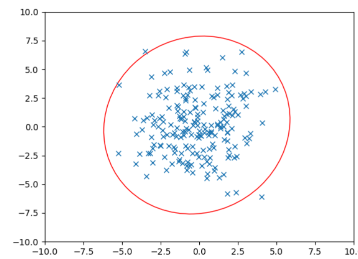
(j) When applying the Central Limit Theorem to approximate the sum of n (independent) Bernoulli random variables with a Gaussian (Normal) distribution, we use a continuity correction because:

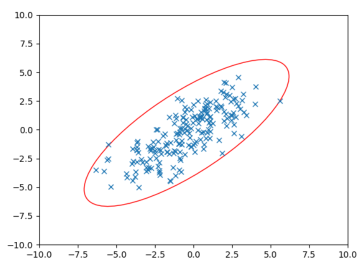
◯ The CLT only applies to a sum of continuous i.i.d random variables.
◯ The continuity correction guarantees we only ever over-estimate the probability of the event we're interested in.
◯ The continuity correction tends to make estimates more accurate, especially for small n.
◯ The continuity correction tends to make computations simpler.

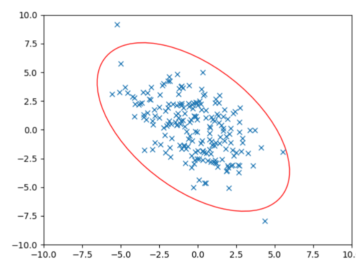
2. Small Problems [11 points]

(a) One of the images below is 200 independent draws from a 2D Gaussian X, Y with mean $\begin{bmatrix}\mathbb{E}[X]\\ \mathbb{E}[Y]\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}$ and covariance matrix $\begin{bmatrix}4 & -3\\ -3 & 7\end{bmatrix}$. Each image shows the range $[-10, 10]$ on both axes.
Fill in the circle next to the correct image. [3 points]









(b) Let $X \sim \mathcal{N}(1, 4)$. Give $\mathbb{P}(X \leq 2)$ as a decimal (you must give a decimal here, you may not use Φ notation in your final answer). [4 points]

(c) A donut shop sells boxes of 12 donuts; they have flavors of chocolate, strawberry, blueberry, and coconut. It's nearing closing time, so there are only two strawberry donuts left (but an unlimited number of the others). How many distinct boxes are possible?
Your answer should be relatively simple; summation symbols or "..." are not simple, but a few plus signs are fine to leave unsimplified. [4 points]

3. [RV Zoo] Wacky Races [25 points]

Robbie and three TAs decide to have a friendly race around Seattle, to decide who will have to do all of the final exam grading. Contestant race in their own wacky vehicles that can finish the race in different amounts of time, independent of each other. Here are the contestants and their wacky vehicles:

- Robbie with his Rubbernoulli-Burner finishes in $R = 10X + 30$ minutes, where $X \sim \text{Ber}(p)$. (Treat p as an unknown constant).
- Charles with his Binomiobile finishes in $B \sim \text{Bin}(128, 0.25)$ minutes.
- Allie with her Gaussoline-Guzzler finishes in $G \sim \mathcal{N}(\mu = 33, \sigma^2 = 9)$ minutes.
- Edward with his Exponentiator 9000 finishes in $E \sim \text{Exp}(\frac{3}{100})$ minutes.

(a) Give the PMF of R. [4 points]

(b) What is the probability that Robbie with his Rubbernoulli-Burner finishes the race before Edward with his Exponentiator 9000, that is, $\mathbb{P}(R < E)$? Use the law of total probability, partitioning on the two possible values of R. Do **not** evaluate or simplify; leave your expression in terms of expressions like "$\mathbb{P}(R = 30)$" (4 points)

(c) What is the probability that Edward with his Exponentiator 9000 takes over 30 minutes to finish the race, that is, $\mathbb{P}(E > 30)$? Your answer must be an expression that could be evaluated with just a scientific calculator (e.g., no integrals). [3 points]

(d) Suppose Robbie with his Rubbernoulli-Burner finished the race before Edward with his Exponentiator 9000. Use Bayes' Theorem to calculate the probability that Robbie finished in 30 minutes, that is, $\mathbb{P}(R = 30 | R < E)$. Your final answer can be given in terms of p. You do not need to fully simplify, but only expressions involving p and constants that can be evaluated with a calculator may remain (still no integrals). [10 points]

(e) Give the expectation of the time for each of the vehicles to finish. Put your final answer directly on the line, in as simplified a form as you can. [4 points total]

- Robbie with his Rubbernoulli-Burner finishes $R = 10X + 30$ minutes, where $X \sim$ Ber(p)

  $\mathbb{E}[R] =$ ____________________

- Charles with his Binomiobile finishes in $B \sim$ Bin$(128, 0.25)$ minutes.

  $\mathbb{E}[B] =$ ____________________

- Allie with her Gaussoline-Guzzler finishes in $G \sim \mathcal{N}(\mu = 33, \sigma^2 = 9)$ minutes.

  $\mathbb{E}[G] =$ ____________________

- Edward with his Exponentiator 9000 finishes in $E \sim$ Exp$(\frac{3}{100})$ minutes.

  $\mathbb{E}[E] =$ ____________________

4. My Cabbages! [20 points]

A cabbage merchant tries to sell cabbages from his cart in Omashu. On typical days, he earns 6 gold pieces from selling his cabbages. However, if the Avatar is in Omashu on a day, the merchant will lose his entire cart of cabbages, causing a loss of 15 gold pieces (and no earnings). The Avatar spends $\frac{1}{3}$ of his time in Omashu, independently choosing each day whether to be in the city.

Let X be the net profit the merchant earns in 20 days.

(a) What is the expected net profit the merchant earns in 20 days, i.e., $\mathbb{E}[X]$? [5 points]

(b) What is Var (X)? [5 points]

The cabbage merchant decides to stock two types of cabbages: red and green. On a particular day, he stocks 10 red cabbages and 4 green cabbages. Every customer randomly chooses one of the **not-yet-purchased** cabbages to buy.

Let R be the event that the **first** cabbage bought is red. Let G be the event that the **second** cabbage bought is green.

(c) What is $\mathbb{P}(G|R)$? [3 points]

(d) What is $\mathbb{P}(G)$? [4 points]

(e) Are R and G independent? Justify your answer in reference to your results from (c) and (d) or a new computation. [3 points]

◯ Yes, they are independent.
◯ No, they are not independent

Put your explanation here.

5. [Concentration] Wild Berries [15 points]

You've stumbled upon the perfect opportunity to mesh two of your favorite hobbies - hiking and baking. There are multiple scenic hikes in Washington that are known for their abundance of berries. In order to bake a blueberry pie, you need at least 260 juicy blueberries.

(a) The first hike you go on, is the Shannon Ridge hike which is near the Mt. Baker area. On average, people have reported collecting 220 blueberries on this seven mile hike. Use Markov's inequality to bound the probability that you collect at least 260 berries on this hike.

(b) After doing more research, you learn that the number of blueberries collected on this hike is distributed with variance 80. Use the Chebyshev's inequality to bound the same probability from part (a). In your answer you must precisely state the event which you are using Chebyshev to bound.

(c) Over the summer months, you plan to go on four more hikes (so five hikes in total), where the number of berries collected in each hike is distributed the same way as the Shannon Ridge hike (but each hike is independent). That is, you will get five independent trails with average 220 berries and variance 80.

Bound the probability that none of your hikes have enough berries for you to make a pie. (Since berries taste best when fresh, all blueberries used in the pie must be collected on the same hike). Write an inequality, and name the tool you used to get it (Hint: use your answer from (b)).

6. [MLE] Swing For The Fences [13 points]

You are a big baseball fan, who loves watching home runs (and therefore hates waiting to watch home runs). Luckily, you are a fan of a very consistent team. Each player has identical hitting skills, which are independent of all their teammates.

If it's a **windy** day, it's easy to hit home runs. Each player hits a home run with probability $3/5$.
If it's a **calm** day, it's hard to hit home runs. Each player hits a home run with probability only $1/5$.

You may assume the only two types of weather are calm and windy.

You watch batters hit, and record how many players hit until the first time one hits a home run. Today, the fourth batter was the first to hit a home run. In this problem, you will use this information to determine the maximum likelihood estimator for the state of the weather (i.e., whether it was windy or calm). In this problem, you are estimating a Boolean parameter ($\{\texttt{windy}, \texttt{calm}\}$) which is different than what we've done in examples from class.

(a) What is the likelihood of it taking (exactly) 4 hitters to see the first home run on a windy day? You do not have to simplify. [4 points]

(b) What is the likelihood of it taking (exactly) 4 hitters to see the first home run on a calm day? You do not have to simplify. [4 points]

(c) After observing it take (exactly) 4 hitters to see the first home run, describe the process you would use to determine the MLE if you had access to a calculator (and/or computer). You should describe in enough detail that someone who doesn't know what an MLE is (but does have a calculator and computer) could perform the computation for you. [3 points]

(d) Parts a and b sound a lot like a Bayes' Rule problem! You don't have enough information to use Bayes' Rule here. What information are you missing that you would need to find $\mathbb{P}(\text{windy}|\text{exactly } 4 \text{ hitters})$? You may give your answer in notation or in English, but if you use notation be sure we'll understand any variables or events you refer to. [2 points]

7. Grading Morale [1 point]

This time, it's **your** turn to give us a probability problem. Come up with your own probability question for us to answer; the wackier, the better. (**As long as you don't leave this page blank, you'll get full credit for this part.**)

Use this page for extra space if you need it. Be sure to tell us to look here.

Use this page for *even more* extra space if you need it. Be sure to tell us to look here.

Reference Sheet: Counting, Discrete Probability

Theorem: Binomial Theorem

Let $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$ a positive integer. Then: $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.

Theorem: Principle of Inclusion-Exclusion (PIE)

2 events: $|A \cup B| = |A| + |B| - |A \cap B|$
3 events: $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$
k events: singles - doubles + triples - quads + ...

Theorem: Pigeonhole Principle

If there are n pigeons we want to put into k holes (where $n > k$), then at least one pigeonhole must contain at least 2 (or to be precise, $\lceil n/k \rceil$) pigeons.

Definition: Key Probability Definitions

The **sample space** is the set Ω of all possible outcomes of an experiment.
An **event** is any subset $E \subseteq \Omega$.
Events E and F are **mutually exclusive** if $E \cap F = \emptyset$.

Definition: Probability space

A *probablity space* is a pair $(\Omega, \mathbb{P})$, where Ω is the sample space
$\mathbb{P} : \Omega \to [0,1]$ is a *probability measure* such that $\sum_{x \in \Omega} \mathbb{P}(x) = 1$.
The probability of an event $E \subseteq \Omega$ is $\mathbb{P}(E) = \sum_{x \in E} \mathbb{P}(x)$.

Definition: Conditional Probability

$\mathbb{P}[A \mid B] = \dfrac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$

Theorem: Bayes Theorem

$\mathbb{P}[A \mid B] = \dfrac{\mathbb{P}[B \mid A]\,\mathbb{P}[A]}{\mathbb{P}[B]}$

Definition: Partition

Non-empty events $E_1, \ldots, E_n$ **partition** the sample space Ω if:

- **(Exhaustive)** $E_1 \cup E_2 \cup \cdots \cup E_n = \bigcup_{i=1}^{n} E_i = \Omega$ (they cover the entire sample space).
- **(Pairwise Mutually Exclusive)** For all $i \neq j$, $E_i \cap E_j = \emptyset$ (none of them overlap)

Theorem: Law of Total Probability (LTP)

If events $E_1, \ldots, E_n$ partition Ω, then for any event F:

$$\mathbb{P}[F] = \sum_{i=1}^{n} \mathbb{P}[F \cap E_i] = \sum_{i=1}^{n} \mathbb{P}[F \mid E_i]\,\mathbb{P}[E_i]$$

Theorem: Bayes Theorem with LTP

Let events $E_1, \ldots, E_n$ partition the sample space Ω, and let F be another event. Then:

$$\mathbb{P}[E_1 \mid F] = \frac{\mathbb{P}[F \mid E_1]\,\mathbb{P}[E_1]}{\sum_{i=1}^{n} \mathbb{P}[F \mid E_i]\,\mathbb{P}[E_i]}$$

Definition: Independence (Events)

A and B are **independent** if any of the following equivalent statements hold:
1. $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]$
2. $\mathbb{P}[A \mid B] = \mathbb{P}[A]$
3. $\mathbb{P}[B \mid A] = \mathbb{P}[B]$

Theorem: Chain Rule

Let $A_1, \ldots, A_n$ be events with nonzero probabilities. Then:
$\mathbb{P}[A_1 \cap \cdots \cap A_n] =$
$\mathbb{P}[A_1]\,\mathbb{P}[A_2 \mid A_1]\,\mathbb{P}[A_3 \mid A_1 \cap A_2] \cdots \mathbb{P}[A_n \mid A_1 \cap \cdots \cap A_{n-1}]$

Definition: Mutual Independence (Events)

We say n events $A_1, A_2, \ldots, A_n$ are **(mutually) independent** if, for *any* subset $I \subseteq [n] = \{1, 2, \ldots, n\}$, we have

$$\mathbb{P}\left[\bigcap_{i \in I} A_i\right] = \prod_{i \in I} \mathbb{P}[A_i]$$

This equation is actually representing 2^n equations since there are 2^n subsets of $[n]$.

Definition: Conditional Independence

A and B are **conditionally independent given an event** C if any of the following equivalent statements hold:
1. $\mathbb{P}[A \cap B \mid C] = \mathbb{P}[A \mid C]\,\mathbb{P}[B \mid C]$
2. $\mathbb{P}[A \mid B \cap C] = \mathbb{P}[A \mid C]$
3. $\mathbb{P}[B \mid A \cap C] = \mathbb{P}[B \mid C]$

Definition: Random Variable (RV)

A random variable X is a function of the outcome $X : \Omega \to \mathbb{R}$. The set of possible values X can take on is its **range/support**, denoted Ω_X.

Definition: Probability Mass Function (PMF)

For a discrete RV X, assigns probabilities to values in its range. That is $p_X : \Omega_X \to [0,1]$ where: $p_X(k) = \mathbb{P}[X = k]$.

Definition: Expectation

The **expectation** of a discrete RV X is: $\mathbb{E}[X] = \sum_{k \in \Omega_X} k \cdot p_X(k)$.

Theorem: Linearity of Expectation (LoE)

For any random variables X, Y (possibly dependent):
$\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$

Theorem: Law of the Unconscious Statistician (LOTUS)

For a discrete RV X and function g, $\mathbb{E}[g(X)] = \sum_{b \in \Omega_X} g(b) \cdot p_X(b)$.

Definition: Variance

$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$.

Theorem: Property of Variance

$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$.

Definition: Independence (Random Variables)

Random variables X and Y are **independent** if for *all* $x \in \Omega_X$ and all $y \in \Omega_Y$:
$\mathbb{P}[X = x, Y = y] = \mathbb{P}[X = x] \cdot \mathbb{P}[Y = y]$.

Theorem: Variance Adds for Independent RVs

If X, Y are independent, then $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Definition: Standard Deviation (SD)

$\sigma_X = \sqrt{\mathrm{Var}(X)}$.

Reference: Continuous and Multivariate Probability

Definition: Cumulative Distribution Function (CDF)

The **cumulative distribution function (CDF)** of ANY random variable is $F_X(t) = \mathbb{P}[X \le t]$.
If X is a *continuous* RV, $F_X(t) = \mathbb{P}[X \le t] = \int_{-\infty}^{t} f_X(w)\, dw$.

Theorem: Multiplicativity of expectation

For any independent random variables X, Y:
$$\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

Definition: Expectation (Continuous)

The **expectation** of a continuous RV X is:
$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f_X(x)\, dx.$$

Theorem: Law of the Unconscious Statistician (LOTUS)

For a continuous RV X: $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$.

Definition: Independent and Identically Distributed (i.i.d.)

We say $X_1, \ldots, X_n$ are said to be **independent and identically distributed (i.i.d.)** if all the X_i's are independent of each other, and have the same distribution (PMF for discrete RVs, or CDF for continuous RVs).

Definition: Joint PMFs

The joint PMF of discrete RVs X and Y is:
$$p_{X,Y}(a, b) = \mathbb{P}[X = a, Y = b]$$
Their joint range is
$$\Omega_{X,Y} = \{(c, d) : p_{X,Y}(c, d) > 0\} \subseteq \Omega_X \times \Omega_Y$$
Note that $\sum_{(s,t)\in\Omega_{X,Y}} p_{X,Y}(s, t) = 1$.

Definition: Joint PDFs

The joint PDF of continuous RVs X and Y is:
$$f_{X,Y}(a, b) \ge 0$$
Their joint range is
$$\Omega_{X,Y} = \{(c, d) : f_{X,Y}(c, d) > 0\} \subseteq \Omega_X \times \Omega_Y$$
Note that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(u, v)\,du\,dv = 1$.

Definition: Marginal PMFs

Let X, Y be discrete random variables. The marginal PMF of X is: $p_X(a) = \sum_{b\in\Omega_Y} p_{X,Y}(a, b)$.

Definition: Marginal PDFs

Let X, Y be continuous random variables. The marginal PDF of X is: $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dy$.

Definition: Independence of RVs (Continuous)

Continuous RVs X, Y are independent, written $X \perp Y$, if for all $x \in \Omega_X$ and $y \in \Omega_Y$,
$$f_{X,Y}(x, y) = f_X(x) f_Y(y).$$

Definition: Conditional Expectation

If X is discrete (and Y is either discrete or continuous), then we define the conditional expectation of $g(X)$ given (the event that) $Y = y$ as:
$$\mathbb{E}[\, g(X) \mid Y = y\,] = \sum_{x\in\Omega_X} g(x)\, \mathbb{P}(X = x \mid Y = y)$$
If X is continuous (and Y is either discrete or continuous), then
$$\mathbb{E}[\, g(X) \mid Y = y\,] = \int_{-\infty}^{\infty} g(x)\, \frac{f_{X,Y}(x, y)}{f_Y(y)}\, dx$$

Theorem: Law of Total Expectation (LTE)

Let X, Y be jointly distributed random variables.

If Y is discrete (and X is either discrete or continuous), then:
$$\mathbb{E}[\, g(X)\,] = \sum_{y\in\Omega_Y} \mathbb{E}[\, g(X) \mid Y = y\,]\; p_Y(y)$$
If Y is continuous (and X is either discrete or continuous), then
$$\mathbb{E}[\, g(X)\,] = \int_{-\infty}^{\infty} \mathbb{E}[\, g(X) \mid Y = y\,]\; f_Y(y)\,dy$$

Reference: Tail Bounds

Theorem: Markov's Inequality

Let $X \ge 0$ be a **non-negative** RV, and let $k > 0$. Then:
$$\mathbb{P}[X \ge k] \le \frac{\mathbb{E}[X]}{k}$$
.

Theorem: Chebyshev's Inequality

Let X be any RV with expected value $\mu = \mathbb{E}[X]$ and finite variance $\text{Var}(X)$. Then, for any real number $\alpha > 0$. Then,
$$\mathbb{P}[|X - \mu| \ge \alpha] \le \frac{\text{Var}(X)}{\alpha^2}$$
.

Theorem: Chernoff Bound

Let $X = X_1 + X_2 + \ldots + X_n$, where $X_1, X_2, \ldots, X_n$ are independent random variables, each taking values in $[0, 1]$. Also, let $\mu = \mathbb{E}[X]$.
For any $1 > \delta > 0$:
$$\mathbb{P}(X \ge (1 + \delta)\mu) \le \exp\left(-\delta^2\mu/3\right)$$
$$\mathbb{P}(X \le (1 - \delta)\mu) \le \exp\left(-\delta^2\mu/2\right)$$

Theorem: The Union Bound

Let $E_1, E_2, \ldots, E_n$ be a collection of events. Then:
$$\mathbb{P}\left[\bigcup_{i=1}^{n} E_i\right] \le \sum_{i=1}^{n} \mathbb{P}[E_i]$$

Reference: Zoo

Definition: Bernoulli/Indicator Random Variable

$X \sim$ Bernoulli(p) (Ber(p) for short) iff X has PMF:

$$p_X(k) = \begin{cases} p, & k = 1 \\ 1-p, & k = 0 \end{cases}$$

$\mathbb{E}[X] = p$ and $\text{Var}(X) = p(1-p)$.

Definition: Binomial Random Variable

$X \sim$ Binomial(n, p) ($\mathsf{Bin}(n,p)$ for short) iff X has PMF

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \Omega_X = \{0, 1, \dots, n\}$$

$\mathbb{E}[X] = np$ and $\text{Var}(X) = np(1-p)$.

Definition: Uniform Random Variable (Discrete)

$X \sim$ Uniform(a, b) (Unif(a, b) for short), for integers $a \leq b$, iff X has PMF:

$$p_X(k) = \frac{1}{b-a+1}, \quad k \in \Omega_X = \{a, a+1, \dots, b\}$$

$\mathbb{E}[X] = \frac{a+b}{2}$ and $\text{Var}(X) = \frac{(b-a)(b-a+2)}{12}$.

Definition: Geometric Random Variable

$X \sim$ Geometric(p) (Geo(p) for short) iff X has PMF:

$$p_X(k) = (1-p)^{k-1} p, \quad k \in \Omega_X = \{1, 2, 3, \dots\}$$

$\mathbb{E}[X] = \frac{1}{p}$ and $\text{Var}(X) = \frac{1-p}{p^2}$.

Definition: Poisson Random Variable

$X \sim$ Poisson(λ) (Poi(λ) for short) iff X has PMF:

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \Omega_X = \{0, 1, 2, \dots\}$$

$\mathbb{E}[X] = \lambda$ and $\text{Var}(X) = \lambda$. If $X_1, \dots, X_n$ are independent Poisson RV's, where $X_i \sim$ Poi(λ_i), then $X = X_1 + \dots + X_n \sim$ Poi$(\lambda_1 + \dots + \lambda_n)$.

Definition: Uniform Random Variable (Continuous)

$X \sim$ Uniform(a, b) (Unif(a, b) for short) iff X has PDF:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in \Omega_X = [a, b] \\ 0 & \text{otherwise} \end{cases}$$

$\mathbb{E}[X] = \frac{a+b}{2}$ and $\text{Var}(X) = \frac{(b-a)^2}{12}$.

Definition: Exponential Random Variable

$X \sim$ Exponential(λ) (Exp(λ) for short) iff X has PDF:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \in \Omega_X = [0, \infty) \\ 0 & \text{otherwise} \end{cases}$$

$\mathbb{E}[X] = \frac{1}{\lambda}$ and $\text{Var}(X) = \frac{1}{\lambda^2}$.

$F_X(x) = 1 - e^{-\lambda x}$ for $x \geq 0$.

Definition: Normal (Gaussian, "bell curve") Random Variable

$X \sim \mathcal{N}(\mu, \sigma^2)$ iff X has PDF:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}, \quad x \in \Omega_X = \mathbb{R}$$

$\mathbb{E}[X] = \mu$ and $\text{Var}(X) = \sigma^2$.

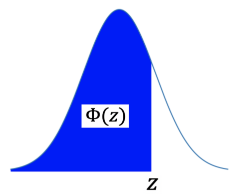
Theorem: Closure of the Normal Under Scale and Shift

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.
In particular, we can always scale/shift to get the standard Normal: $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$.

Theorem: Closure of the Normal Under Addition

If $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ are *independent*, then

$$aX + bY + c \sim \mathcal{N}(a\mu_X + b\mu_Y + c, a^2\sigma_X^2 + b^2\sigma_Y^2)$$



Φ Table: $\mathbb{P}(Z \leq z)$ when $Z \sim \mathcal{N}(0, 1)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.9452	0.9463	0.94738	0.94845	0.9495	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.9608	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.9732	0.97381	0.97441	0.975	0.97558	0.97615	0.9767
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.9803	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.983	0.98341	0.98382	0.98422	0.98461	0.985	0.98537	0.98574
2.2	0.9861	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.9884	0.9887	0.98899
2.3	0.98928	0.98956	0.98983	0.9901	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.9918	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.9943	0.99446	0.99461	0.99477	0.99492	0.99506	0.9952
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.999