CSE 312 Step-by-Step Problem Guide

By Hannah Swoffer

In CSE 312, there are some problem types that have generic steps. This guide compiles the plain English problem strategies from lecture, breaking down *how* to approach specific types of problems using examples.

Many examples reference Claris Winston's 24su lectures.

---

1. Rearranging Elements

> 1. Count the elements
>   - If there are n elements, then there are $n!$ rearrangements.
> 2. *If there are duplicate elements,* divide out over counting
>   - If there are duplicates with m elements, divide out $m!$ arrangements for each set of duplicates

Example

COFFEE

1. There are 6 elements, implying 6! arrangements.
2. There are duplicate letters (**F** and **E**), so we must divide out 2! twice for over counting.

$$\boxed{\frac{6!}{2!2!} = 180 \text{ rearrangements}}$$

2. Stars and Bars Problems

> *If you see language like...*
>
> - Items are indistinguishable
> - How many of each *type* of item? (We don't care about the specific items we pick, just the number of each type)

Use Stars and Bars

1. Find n, the number of indistinguishable items (stars)
2. Find k, the number of distinguishable items (bars)
3. Apply stars and bars

$$\binom{n+k-1}{k-1} = \binom{n+k-1}{n}$$

Example

How many different ways can I distribute 10 indistinguishable candies into 3 containers?

1. $n = 10$ since we have 10 *indistinguishable* candies
2. $k = 3$ since we have 3 *distinguishable* containers
3. Apply stars and bars

$$\boxed{\binom{10+3-1}{3-1} = \binom{12}{2} \text{ ways}}$$

3. Combinatorial Proofs

1. Describe a scenario
2. Explain how the LHS counts the outcomes in that scenario
3. Explain how the RHS counts the outcomes in the same scenario
4. State that "because the LHS and RHS both count the number of outcomes in the same scenario, they must be equal" (this is our ultimate goal)

Example

We want to take some corgis on a walk. Use this scenario to prove:

$$\binom{n}{k} = \binom{n}{n-k}$$

1. Consider n corgis and we want to choose a subset of k to take on a walk.
2. The LHS represents the number of ways to choose k corgis from a group of n corgis. To do this, we choose k corgis from n available, and $\binom{n}{k}$ directly counts this.
3. The RHS represents the number of ways to choose $n-k$ corgis from the same group of n corgis. To find this, we have $\binom{n}{n-k}$ the corgis that will *not* go on the walk, which means the remaining k corgis will go on the walk.
4. Because both the LHS $\binom{n}{k}$ and RHS $\binom{n}{n-k}$ count the number of ways to choose k corgis to go on the walk from the group of n corgis, they must be equal. Therefore,
$$\binom{n}{k} = \binom{n}{n-k}$$

4. Union Problems

If you see language like...

- Either A or B
- At least one of A or B

We are finding the union of some sets, so our goal is to find $|A \cup B|$

1. *If the sets are disjoint*, use the sum rule
2. *Otherwise*, use the Principle of Inclusion-Exclusion
3. *If the above doesn't work*, take the complement and find the intersection

Example

Example: 20 people were surveyed about their music opinions. 7 people liked Ariana Grande, 10 people liked Bruno Mars, and 3 people liked both. How many people liked at least one of the artists?

1. The sets are not disjoint because 3 people like both artists
2. Therefore, we can use the Principle of Inclusion Exclusion

   $A \sim$ number of people that like Ariana Grande
   $B \sim$ number of people that like Bruno Mars
   $|A \cap B| \sim$ number of people that like both

$$|A \cup B| = |A| + |B| - |A \cap B| = \boxed{14 \text{ people}}$$

5. Pigeonhole Principle

1. Define the pigeons
2. Define the pigeonholes
3. Describe how to map from the pigeons to the pigeonholes
4. State: by the pigeonhole principle, there are at least $\lceil \frac{n}{k} \rceil$ pigeons in the same pigeonhole. Therefore, there are at least $\lceil \frac{n}{k} \rceil$...

Example

Are there at least two students in a university who have read the same number of books?

1. Pigeons: all students at the university ($20,000$ students at the university)
2. Pigeonholes: all possible book counts (students can read up to 5,000 books)
3. Mapping: every student is "assigned" to a particular book count
4. By the pigeonhole principle, there are at least $\left\lceil \frac{20{,}000}{5{,}000} \right\rceil = 4$ pigeons in the same pigeonhole. Therefore, there are at least 4 students with the same book count.

6. Uniform Probability Spaces

If you see language like...

- Equally likely
- Randomly choosing
- Uniformly at random

We are finding the uniform probability of an event.

1. Define the sample space: Ω
2. Find the size of the sample space: $|\Omega|$
3. Write the probability measure: $P(\omega) = \frac{1}{|\Omega|} \quad \forall \omega \in \Omega$
4. Define an event: E (for example)
5. Find the size of the event: $|E|$
6. Find the probability of your event: $P(E)$

Example

Example: Suppose you rolled 2 dice (each is fair). What is the probability that both dice add up to 7?

1. $\Omega \sim$ all possible outcomes when rolling two dice
   $\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$
2. $|\Omega| = 6^2 = 36$
3. Each outcome is equally likely, so the probability of any specific outcome (ω) is:
   $$P(\omega) = \frac{1}{36} \quad \forall \omega \in \Omega$$
4. $E \sim$ the event that both dice add up to 7
   $E = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$
5. $|E| = 6$
6. The probability of event E is
   $$P(E) = \frac{|E|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}$$
   Therefore, the probability that the sum of the two dice is 7 is $\boxed{\frac{1}{6}}$.

7. Conditional Probability

If you see language like...

- If
- Given

This indicates that we might need to use conditional probability. The steps below could be in any order or could be used multiple times, depending on the problem's context.

1. Start by defining events, writing down the given probabilities, and determining the probability you are interested in finding.
2. *If you are finding the probability of a conditioned event (ex: $P(A|B)$) and you are told the reverse of such event (ex: $P(B|A)$)*, use Bayes' Theorem.

3. *If you are finding the probability of an event (ex: $P(A)$), and you are told the probability of such event conditioned on other events (ex: $P(A|B), P(A|B^c)$),* use the Law of Total Probability.

4. *If the previous approaches cannot be applied*, try the Definition of Conditional probability.

5. Use any combination of the tools listed above to find your probability.

Example

Suppose there is a famous art museum with two galleries: Gallery A and Gallery B. One gallery displays *The Starry Night* by Vincent van Gogh, while the other gallery displays a very similar replica. A visitor to the museum has seen a painting that they think is the original. The probability that *The Starry Night* is in Gallery A is 0.9, and the probability that *The Starry Night* is in Gallery B is 0.1. The visitor identifies the painting correctly 95% of the time in Gallery A and 10% of the time in Gallery B. Given that the visitor has identified the painting correctly, what is the probability that it is actually *The Starry Night* from Gallery A?

We see the keyword "given," so let's try using conditional probability.

1. Define events and gather probabilities:

   $A \sim$ the event that the painting is from Gallery A
   $B \sim$ the event that the painting is from Gallery B
   $C \sim$ the event that the painting is *The Starry Night*

   $P(A) = 0.9$
   $P(B) = 0.1$
   $P(C|A) = 0.95$
   $P(C|B) = 0.1$
   $P(A|C) = ?$

   Continued on the next page...

Example

2. Since we are trying to find the reverse of information we are already given, this indicates we should use Baye's Theorem.

$$\begin{aligned} P(A|C) &= \frac{P(C|A) \cdot P(A)}{P(C)} \\ &= \frac{0.95 \cdot 0.9}{P(C)} \end{aligned}$$

3. $P(C)$ is unknown, but we do have information about the probability of C conditioned on another event. Therefore, we can use the Law of Total Probability. Since there are only 2 galleries, we only need to consider Gallery A and Gallery B to find the total probability.

$$\begin{aligned} P(C) &= P(C|A) \cdot P(A) + P(C|B) \cdot P(B) \\ &= 0.95 \cdot 0.9 + 0.1 \cdot 0.1 \\ &= 0.865 \end{aligned}$$

4. No other conditional probability is required, as we have already found our needed information.

5. Now we can use $P(C)$ to complete our calculation with Baye's Theorem:

$$P(A|C) = \frac{0.95 \cdot 0.9}{0.865} \approx 0.9884$$

Therefore, the probability that the painting is actually *The Starry Night* from Gallery A, given that it is identified as *The Starry Night*, is approximately $\boxed{98.84\%}$.

8. Independence

1. Gather information

2. Check if $P(A \cap B) = P(A) \cdot P(B)$

OR

Check if $P(A|B) = P(A)$

Example

We roll a fair 6 sided die twice. Let A be the event that we roll a number greater than 4 on the first roll (i.e. rolling a 5 or 6). Let B be the event that we roll a number less than or equal to 3 on the second roll (i.e. 1, 2, or 3). Let $P(A \cap B) = \frac{1}{12}$. Find $P(A)$ and $P(B)$ and check if they are independent events.

1. Gather information

$$\begin{aligned} P(A \cap B) &= \frac{1}{12} && \text{given} \\ P(A) &= \frac{2}{6} = \frac{1}{3} && \text{divide rolling a 5 or 6 by the total} \\ P(B) &= \frac{3}{6} = \frac{1}{2} && \text{divide rolling a 1, 2, or 3 by the total} \\ P(A) \cdot P(B) &= \frac{1}{6} \end{aligned}$$

2. Check Independence

$$P(A \cap B) = P(A) \cdot P(B)$$
$$\frac{1}{12} \neq \frac{1}{6}$$

So the events A and B are *not* independent.

Example

Building intuition with independence:

After we flip a coin and pick which dice to roll, we roll that dice twice independently

- The outcomes of the two dice rolls are conditionally independent on the coin flip
- After we know the outcome of the coin flip and which dice we're rolling, the dice is rolled independently

Given the patient has a disease, Test A and Test B correctly detect the disease independently

- Both tests are conditionally independent on the event the patient has the disease

Each customer chooses to buy a latte independently

- The events for each customer buying a latte are mutually independent

9. Strategies For Finding Expectation

1. *If the support is small enough and we can find the PMF from X*, we can use the Definition of Expectation.

$$E[X] = \sum_{k \in \Omega_k} k \cdot P(X = k)$$

2. *If we can break X into a sum of indicator random variables*, we can use Linearity of Expectation.

$$E[X + Y] = E[X] + E[Y]$$

3. *If its a function of X*, we can use the Law of the Unconscious Statistician (LOTUS).

$$E[g(X)] = \sum_{k \in \Omega_k} g(k) \cdot P(X = k)$$

4. *If a random variable follows the distribution from one of the random variables in the zoo*, then we can determine the parameters and plug into the expectation formula.

5. *If the random variable depends on some other events that partition the sample space, and computing the expectation given each of those events we're partitioning on is easier than directly finding* $E[X]$, use the Law of Total Expectation.

A common strategy for tackling challenging expectation problems where we use linearity of expectation is:

1. Decompose: Decompose the random variable into a sum of simpler random variables
$$X = \sum_{i=1}^{n} X_n$$
2. LOE: Apply Linearity of Expectation
$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$
or
$$E[X] = E[X_1] + E[X_2] + ... + E[X_n]$$
3. Conquer: Compute the expectation of each $E[X_i]$ (where X_i is often an indicator random variable)

Example

At a quaint restaurant in Italy, there are n different types of pizza. Each of m friends orders one pizza, and each friend chooses a pizza uniformly at random. We want to find the expected number of different pizzas ordered by the group.

1. Decompose:
   $X \sim$ total types of pizza ordered by at least one friend in the group
   $X_i \sim$ the i^{th} pizza ordered by at least one friend in the group
   $X_i = \begin{cases} 1 & \text{if at least one friend orders pizza type } i \\ 0 & \text{otherwise} \end{cases}$

$$X = \sum_{i=1}^{n} X_i$$

2. LOE:

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

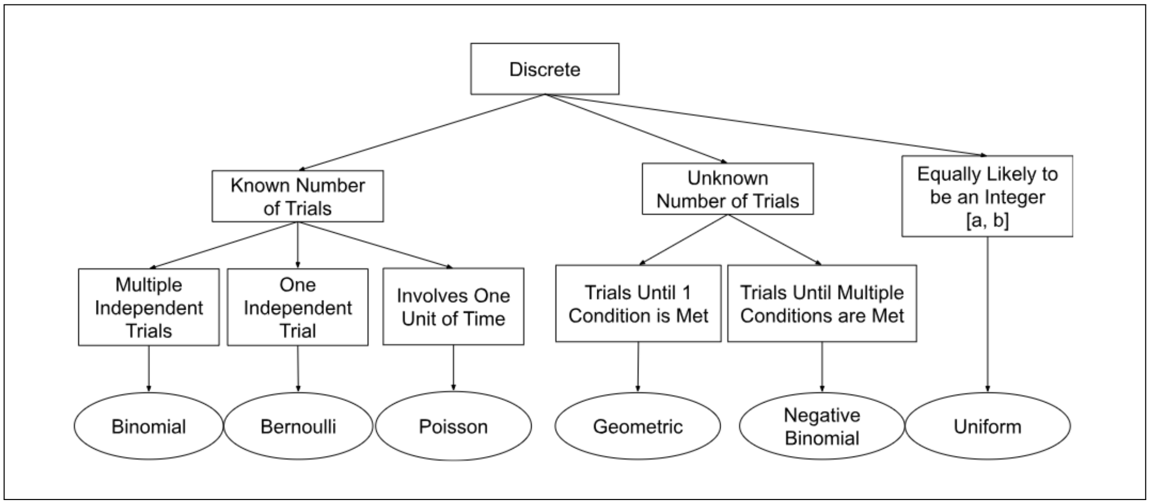
3. Conquer:

   - To compute $\mathbb{E}[X_i]$, we first find the probability that pizza type i is ordered by at least one friend, which is $\frac{1}{n}$.
   - We want to use complementary counting here because of the keywords *at least*. The probability that pizza type i is not ordered by a single friend is $1 - \frac{1}{n}$. Since each friend orders independently, the probability that pizza type i is not ordered by any of the m friends is $\left(1 - \frac{1}{n}\right)^m$.
   - Therefore, the probability that pizza type i is ordered by at least one friend is:
   $$1 - \left(1 - \frac{1}{n}\right)^m$$
   - Therefore, the expectation $\mathbb{E}[X_i]$ is:
   $$E[X_i] = 1 - \left(1 - \frac{1}{n}\right)^m$$

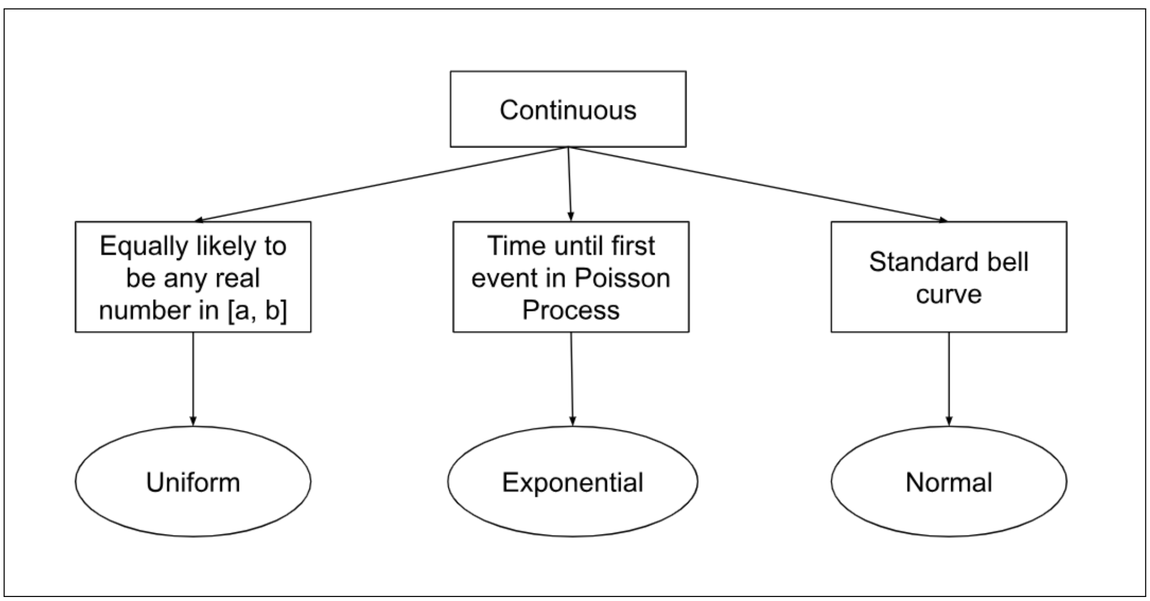
   Returning to our LOE and applying this to all n pizza types, we get:

$$\boxed{E[X] = n\left(1 - \left(1 - \frac{1}{n}\right)^m\right)}$$

10. Identifying Discrete Distributions



11. Identifying Continuous Distributions



12. Central Limit Theorem

1. Setup the problem and write the event you are interested in, in terms of sum of random variables. We can ask ourselves - what do we want to solve for and what is the probability we want to be true?

- Write the random variable we're interested in as a sum of i.i.d., random variables
- Apply CLT to $X = X_1 + X_2 + ... + X_n$ (we can approximate X as a normal random variable $Y \sim N(n\mu, n\sigma^2)$)
- Write the probability we are interested in

2. *If the random variables are discrete,* apply continuity correction
   - Write down the discrete values in the support of X
   - Look at the difference between those values to find the "bucket" we would assign each of those values to
   - Think about what values of X would correspond to the probability we're interested in, and "correct" the interval so we are looking at the entire continuous interval the values correspond to.
3. *(From here, we're doing a typical probability problem with the normal distribution)* Normalize RV to have mean 0 and standard deviation 1: $Z = \frac{Y-\mu}{\sigma}$
4. Replace RV in the probability expression with $Z \sim N(0, 1)$
5. Write in terms of $\Phi(z) = P(Z \leq x)$
6. Look up in the Phi table (or a reverse Phi table lookup if we're for a value of z that gives a certain probability)

Example

A boat makes 30 separate trips in a month. Each day, the number of nautical miles traveled is an integer between 50 and 100, uniformly distributed. Assume each day's distance is independent of the others. Using this information, estimate the probability that the boat travels between 2,200 and 2,400 nautical miles in a particular month using CLT. Apply continuity correction if necessary.

1. Problem Setup

   - Write the RV as a sum of i.i.d. RVs

     $X \sim$ the total number of nautical miles traveled by a boat in the month (30 trips)
     $X_i \sim$ Unif(50, 100), the number of nautical miles traveled on the i^{th} day of the month

     Since the boat makes 30 trips in a month, the total number of nautical miles traveled in the month can be written as the sum of 30 i.i.d. random variables where each X_i is uniformly distributed between 50 and 100 nautical miles:

     $$X = X_1 + X_2 + ... + X_{30}$$

   - Apply CLT to $X = X_1 + X_2 + ... + X_{30}$ (we can approximate X as a normal random variable $Y \sim N(30 \cdot \mu, 30 \cdot \sigma^2)$)

     To approximate X using the Central Limit Theorem (CLT), we need to find the mean and variance of each X_i

     $$\mu = \frac{50 + 100}{2} = 75$$

     $$\sigma^2 = \frac{(100 - 50)^2}{12} = 625$$

     By the CLT, X can be approximated as

     $$Y \sim N(30 \cdot 75, 30 \cdot 625) = N(2250, 18750)$$

   - Write the probability we are interested in

     $$P(2,200 \leq Y \leq 2,400)$$

Continued on the next page...

Example

2. Since the RVs are discrete, apply continuity correction

$$P(2,200 \le Y \le 2,400) \approx P(2,199.5 \le Y \le 2,400.5)$$

3. Normalize RV to have mean 0 and standard deviation 1: $Z = \frac{Y-\mu}{\sigma}$

$$Z = \frac{Y - 2,250}{\sqrt{18750}} \approx \frac{Y - 2,250}{136.9}$$

4. Replace RV in probability expression with $Z \sim N(0,1)$

$$\begin{aligned}
&P(2,199.5 \le Y \le 2,400.5) \\
&= P\left(\frac{2,199.5 - 2,250}{136.9} \le \frac{Y - 2,250}{136.9} \le \frac{2,400.5 - 2,250}{136.9}\right) \\
&= P\left(-0.37 \le Z \le 1.10\right)
\end{aligned}$$

5. Write in terms of $\Phi(z)$

$$\begin{aligned}
P\left(-0.37 \le Z \le 1.10\right) &= \Phi(1.10) - \Phi(-0.37) \\
&= \Phi(1.10) - [1 - \Phi(0.37)] \\
&= \Phi(1.10) + \Phi(0.37) - 1
\end{aligned}$$

6. Look up in Phi Table

$$\begin{aligned}
\Phi(1.10) + \Phi(0.37) - 1 &= .8643 + .6443 - 1 \\
&= \boxed{0.5086}
\end{aligned}$$

So, the probability that the boat travels between 2,200 and 2,400 nautical miles in a month is approximately 0.5086, or 50.86%.

13. Joint Continuous Probabilities

1. Identify the distribution
2. Write the joint density function
3. Find the probability

Example

Consider a population of raccoons in a forest. Each raccoon spends a certain amount of time foraging for food each day. The time spent foraging by Raccoon A follows an exponential distribution with a mean of 2 hours, and the time spent foraging by Raccoon B follows an exponential distribution with a mean of 3 hours. The two raccoons forage independently of each other. Let X be the time spent foraging by Raccoon A and Y be the time spent foraging by Raccoon B. What is the probability that Raccoon A spends less time foraging than Raccoon B?

1. Identify the distribution
   X follows an exponential distribution with mean 2. Therefore,
   $X \sim \text{Exp}\left(\frac{1}{2}\right)$
   Y follows an exponential distribution with mean 3. Therefore,
   $Y \sim \text{Exp}\left(\frac{1}{3}\right)$

2. Write the joint density function
   Since X and Y are independent:

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$$

$$f_X(x) = \frac{1}{2}e^{-\frac{1}{2}x} \quad \text{for all } x \geq 0$$

$$f_Y(y) = \frac{1}{3}e^{-\frac{1}{3}y} \quad \text{for all } y \geq 0$$

$$f_{X,Y}(x,y) = \left(\frac{1}{2}e^{-\frac{1}{2}x}\right) \cdot \left(\frac{1}{3}e^{-\frac{1}{3}y}\right)$$

3. Find $P(X < Y)$

$$\begin{aligned}
P(X < Y) &= \int_0^\infty \int_0^y f_{X,Y}(x,y)\ dxdy \\
&= \int_0^\infty \int_0^y \left(\frac{1}{2}e^{-\frac{1}{2}x}\right) \cdot \left(\frac{1}{3}e^{-\frac{1}{3}y}\right) dxdy \\
&= \frac{1}{6}\int_0^\infty e^{-\frac{1}{3}y}\left(\int_0^y e^{-\frac{1}{2}x}dx\right) dy \\
&= \frac{1}{6}\int_0^\infty e^{-\frac{1}{3}y} \cdot 2\left(1 - e^{-\frac{1}{2}y}\right) dy \\
&= \frac{1}{3}\int_0^\infty e^{-\frac{1}{3}y}dy - \frac{1}{3}\int_0^\infty e^{-\left(\frac{1}{2}+\frac{1}{3}\right)y}dy \\
&= \boxed{\frac{3}{5}}
\end{aligned}$$

14. Tail Bounds

1. *If X is non-negative and E[X] is known,* use Markov's Inequality
2. *If E[X] and Var(X) are known,* use Chebyshev's Inequality
3. *If X is a sum of independent Bernoulli's,* use a Chernoff Bounf
4. *If we don't have enough information to find the union,* use a Union Bound

Example

In the TV show *Bridgerton*, the amount of time characters spend attending social events (e.g., balls, garden parties) each week is tracked. The mean number of hours spent at these events is 10 hours per week, and the standard deviation is 2 hours. Put an upper bound on $P(X \geq 13)$, where X is the number of hours spent on leisure activities per week.

Since we are given the expected value and can find the variance from the standard deviation, we can use Chebyshev's Inequality.

$$P(|X - E[X]| \geq \alpha) \leq \frac{Var(X)}{\alpha^2}$$

$$\begin{aligned} P(X \geq 13) &= P(X - 10 \geq 13 - 10) \\ &= P(X - 10 \geq 3) \\ &\leq P(|X - 10| \geq 3) && \text{abs. value expands our range} \\ &\leq \frac{Var(X)}{3^2} \\ &\leq \boxed{\frac{4}{9}} \end{aligned}$$

Thus, the probability that a character spends 13 or more hours at social events is at most $\frac{4}{9}$.

15. Maximum Likelihood Estimators

1. Write the likelihood function
2. Take the ln of the likelihood function

3. Take the derivative of the log-likelihood function
4. Set the derivative to 0, and solve for the MLE $\hat{\theta}$
5. Verify it is a maximum with second derivative test (not in 312)

Example

A sunglasses manufacturer classifies defects into three categories:

- Minor defects with probability θ
- Moderate defects with probability 2θ
- Major defects with probability $1 - 3\theta$

In a batch of 60 sunglasses, there are 15 with minor defects, 30 with moderate defects, and 15 with major defects. Find MLE for θ.

1. Likelihood function:

$$\mathcal{L}(x_1, ..., x_{60}; \theta) = \theta^{15} \cdot (2\theta)^{30} \cdot (1 - 3\theta)^{15}$$

2. Log-likelihood function:

$$\begin{aligned} \ln \mathcal{L}(x_1, ..., x_{60}; \theta) &= 15 \ln(\theta) + 30 \ln(2\theta) + 15 \ln(1 - 3\theta) \\ &= 15 \ln(\theta) + 30 \ln(2) + 30 \ln(\theta) + 15 \ln(1 - 3\theta) \\ &= 45 \ln(\theta) + 30 \ln(2) + 15 \ln(1 - 3\theta) \end{aligned}$$

3. Derivative with respect to θ:

$$\begin{aligned} \frac{\partial}{\partial \theta} \ln \mathcal{L}(x_1, ..., x_{60}; \theta) &= \frac{\partial}{\partial \theta}(45 \ln(\theta) + 30 \ln(2) + 15 \ln(1 - 3\theta)) \\ &= \frac{45}{\theta} - \frac{45}{1 - 3\theta} \end{aligned}$$

4. Set equal to 0 and solve for $\hat{\theta}$:

$$\frac{45}{\hat{\theta}} - \frac{45}{1 - 3\hat{\theta}} = 0$$

$$\boxed{\hat{\theta} = \frac{45}{180} = \frac{1}{4}}$$