

MAT 2377 Cheat Sheet

1 Chapter 1: Probabilities

The **sample space** is the set of all possible outcomes.

An **event** is a collection of outcomes in the sample space. Usually this is what we are looking to work with.

We can count items using the k stage procedure.

If we have k stages, each with n_1, n_2, n_3, \dots possibilities, then the total number of possibilities is just $n_1 \cdot n_2 \cdot n_3 \cdot \dots \cdot n_k$.

1.1 Ordered Samples

If we have an ordered sample, then we see that picking 1, 2, 3 is different than picking in a different order 1, 3, 2.

If we draw r items from a bag of n items:

- If we replace each item after drawing, we have: $n \cdot n \cdot n \cdot \dots = n^r$ possibilities
- If we do NOT replace the items, we have: $n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot (n-r) = \frac{n!}{(n-r)!} = {}_n P_r$

1.2 Unordered Samples

This is when the order of the samples does not matter, so 1, 2, 3 would be the same as 1, 3, 2.

We can see the number of unordered samples possible with r draws in a sample space of size n using:

$$\frac{n!}{(n-r)!r!} = {}_n C_r$$

1.3 Probabilities

The probability of an event A with N total outcomes and a favourable outcomes is just:

$$P(A) = \frac{a}{N}$$

We can add probabilities using the following formula:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Any 2 events that satisfy the following expression are called **independent**.

$$P(A \cap B) = P(A) \cdot P(B)$$

1.4 Conditional Probability

We say that the probability of event B given that event A has already happened is:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

1.5 Law of Total Probability

This basically works off of the fact that all probabilities must add up to 1.

This is the specific case to 2 events A and B :

$$P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A})$$

This uses the fact that A and \bar{A} are mutually exclusive, and exhaustive (covers all of S).

So in general, if we have A_1, A_2, \dots, A_k and A_1, A_2, \dots, A_k are mutually exclusive and exhaustive, then we say:

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \dots + P(B|A_k)P(A_k)$$

1.6 Bayes Theorem

This is a way to get the opposite conditional probability to what we have.

If we have $P(A|B)$, among a couple other things, we can obtain $P(B|A)$ with:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

2 Chapter 2: Discrete Random Variables

3 Chapter 3: Continuous Random Variables

4 Chapter 4: Descriptive Statistics and Sampling

To describe a dataset, we have measures of central tendency such as mean and median, and measures of spread such as standard deviation, quartiles, and inter-quartile-range.

The median is just the middle value of a **sorted** dataset. If there are 2 middle values, we just take the mean of both of those.

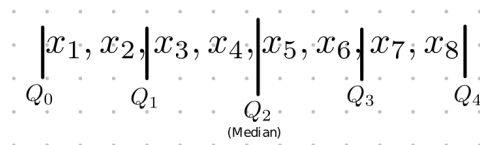
The mean is just:

$$MEAN = \bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

The median is often used since it is not heavily influenced by outliers unlike mean.

4.1 Quartiles

The quartile is like taking the median of the lower half of the data (under the true median).



We call the **Inter Quartile Range (IQR)** as the difference between the third and first quartile $IQR = Q_3 - Q_1$

We identify a datapoint x as an outlier if:

$$x < Q_1 - 1.5IQR \quad \text{or} \quad x > Q_3 + 1.5IQR$$

4.2 Sample Statistics

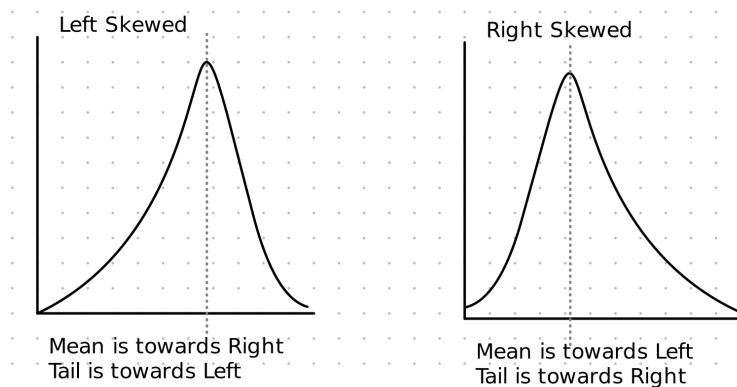
If we do not know the variance of a whole dataset (such as the entire earth's population) then we can consider a sample of this population to estimate the population.

We have **sample standard deviation** s and **sample variance** s^2 . These estimate the standard deviation σ and variance σ^2 respectively.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

4.3 Skewness

We call a dataset **left skewed** if the tail of the data is to the left (outliers on the left) or **right skewed** if the tail of the data is on the right (outliers on the right).



4.4 Independent and Identically Distributed (IID) Case

When all variables are independent and identically distributed we say the expected value and variance of the entire set is just the number of variables times the variance/expected value of one item.

$$\mathbb{E}\left[\sum_{i=1}^n X_i\right] = n\mu \quad \text{Var}\left[\sum_{i=1}^n X_i\right] = n\sigma^2$$

Then we say that if we are considering a sample of these, we have:

$$\mathbb{E}[\bar{X}] = \mu \quad \text{Var}[\bar{X}] = \frac{\sigma^2}{n}$$

We can also use the normal distribution if the population is normally distributed to model $\sum_{i=1}^n X_i$ or \bar{X} .

4.5 Central Limit Theorem

This states that as the number of runs of an experiment, it will start to reach a normal distribution. This is regardless of whether or not the experiment is normal or not.

4.6 Difference between 2 Means

We can work with 2 variables X_1, X_2, \dots, X_n with μ_1, σ_1^2 , and Y_1, Y_2, \dots, Y_m with μ_2, σ_2^2 using the following formula:

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$$

4.7 Other Distributions

We have 2 other main distributions. The Chi squared (χ^2) distribution, and Student's t distribution. Student's t distribution is used when the population variance is unknown, and we have to approximate using the sample variance (standard deviation).

Both of these distributions have the degrees of freedom which is just $n - 1$.

5 Chapter 5: Point and Interval Estimation

This chapter is mostly about confidence intervals (CI).

We have 2 main confidence intervals that we use. Each of them has an α value where if we say the n percent interval, we have $\alpha = 1 - n$.

So for the 95% confidence interval, $\alpha = 0.05$.

The 2 main confidence intervals are the 95 percent, and 99 percent.

5.1 CI When σ is known

We can use a normal distribution to model this since we know σ , n , and \bar{X} .

We use the equation:

$$CI = \bar{X} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Using the normal table we can get $Z_{\alpha/2}$. For $\alpha = 0.05$ we have $Z_{0.025} = 1.96$ and for $\alpha = 0.01$ we have $Z_{0.005} = 2.575$.

5.2 CI When σ is unknown

Here we have to find the sample variance s and we know n , and \bar{X} .

We use Student's t distribution with the equation:

$$CI = \bar{X} \pm t_{\alpha/2}(n-1) \frac{s}{\sqrt{n}}$$

Recall that $n - 1$ is the degrees of freedom for the t distribution.

5.3 CI For a Proportion

When we are dealing with a proportion for a binomial distributions (2 options, either success or failure), we say that P is the probability of success.

We can model this using the normal distribution using:

$$CI = P \pm z_{\alpha/2} \sqrt{\frac{P(1-P)}{n}}$$