First Exam Study Guide

Introduction This summary is not meant to be comprehensive. By using this document, you acknowledge you have read the disclaimer.

Intrepretations Look back at Exam One and Exam Two for the interpretations of things like p-values and other vocabulary. Remember to be concise when presenting these on an exam - sometimes saying more than necessary can hurt more than it helps.

The Formula Card Knowing how to quickly utilize the formula card is helpful for many of the problems you may encounter. The notation on the formula card is correct, and a good habit is to always check the formula card whenever you feel you have hit a dead-end. Often, a formula or hint on the card can provide the next step to finish the problem.

Name that scenerio Being able to identify which test you are looking at quickly and accurately is invaluable. The assumptions, interpretations, distributions, and conclusions all hinge on selecting the correct test, so be sure to study name-that-scenario questions

Broad Concepts The final will test your understanding of broad concepts in this course. Make sure you have a grasp of the big picture as well as the methodology behind each test.

Remember To get a good nights sleep and practice taking an exam towards the later part of the day. This is a later exam than most. Be sure to prepare yourself by taking a practice exam around the time of our actual exam.

General Probability Questions Probability is one of the more difficult units in the class because it sometimes is difficult to decipher what the question is asking. Remember a few simple rules before you start any probability question to start off on the right foot:

1. Most problems involve using one or at most two formulas on the formula card. Begin by defining terms. For example, if we are looking at the probability of it raining vs the probability of you smiling, then let A = "It is raining" and let B = "You are Smiling". Then we can use the formulas on our formula card directly.
2. When you see something that looks *like* coin flips then you are probably dealing with a binomial problem. For instance, if we are trying to see what the probability of 10 out of our 15 neighbors love chocolate, and we know that the probability of loving chocolate is 90%, then this is probably a binomial problem. For this we need:
   - It has a yes/no style question
   - It has a fixed number of "flips"

Mutual Exclusivity and Independence These two are *not* the same thing! The exams generally test this

Mutual Exclusivity This says that the two events cannot happen at the same time. If so $P(A \text{ and } B) = 0$. Example: If we flip a coin, we cannot get a head and a tail in the same toss

Independence This says that knowing something about A will tell us nothing about B. The formula on the formula card is: $P(A|B) = P(A)$, but an easier formula for these problems is to check does this hold: $P(A \text{ and } B) = P(A) \times P(B)$. This formula is not on your formula card but always comes in handy. Example: If I flip a head on turn one and the probability of flipping a head is 50%, the probability that I flip a head on turn two is still 50% on turn two!

Something *cannot* be both mutually exclusive and independent!

Probability Example Let's imagine that we saw the following on the exam:

> The probability of it raining is 30%. The probability of you not smiling is 20%. The probability of it raining or you smiling is 50%.

Problem Setup First thing's first: always set up the problem by putting labels to things! We will let A = Probability of it Raining and B = Probability of you smiling

Writing out the problem in Symbols This step makes things a lot easier since we can now use our formula card

$$P(A) = 30\% \text{ Given in the problem}$$
$$P(\text{Not } B) = 20\% \text{ We have the probabilty of not smiling}$$
$$P(A \text{ or } B) = 50\%$$

Solving for things we want Let's find the probability of smiling

$$P(B) = 1 - P(\text{Not } B) = 1 - 20\% = 80\%$$

Are Smiling and Raining Mutually Exclusive? This is where writing out the problem in symbols helps out a lot because we can use what we have from the formula card:

$$\underbrace{P(A \text{ or } B)}_{50\%} = \underbrace{P(A)}_{30\%} + \underbrace{P(B)}_{80\%} - \underbrace{P(A \text{ and } B)}_{???}$$

and solving out the algebra gives us:

$$P(A \text{ and } B) = P(A \text{ or } B) - P(A) - P(B)$$
$$= 60\%$$

Because $P(A \text{ and } B)$ is not equal to zero, we can say that Smiling and Raining are not mutually exclusive!

Are Smiling and Raining Independent? We can use the fact that Smiling and Raining are Independent only if $P(A) \times P(B) = P(A \text{ and } B)$. Does this hold? We will use the answer from the last problem to help us.

$$\underbrace{P(A)}_{30\%} \times \underbrace{P(B)}_{80\%} \stackrel{?}{=} \underbrace{P(A \text{ and } B)}_{60\%}$$
$$24\% \neq 60\%$$

and thus we can see that A and B (Raining and Smiling) are not independent because the formula is not equal on both sides.

Binomial Example Lets go over how a binomial problem shows up on the exam.

> You are watching CSPAN because you are into politics and you find out (to your surprise) that only 3% of Americans watch CSPAN. You're wondering what the probability that at least one of your neighbors watches CSPAN so you can watch it with them. Assume you have 25 neighbors.

Why is this a Binomial Problem? 1. Binary Choice - Each neighbor either watches CSPAN or does not. These are the only two choices

2. Fixed p - We assume that the probability of one neighbor watching CSPAN is the same as any other neighbor
3. Fixed n - We have a total of 25 neighbors and this doesn't change

What is the probability that if you have one neighbor they watch CSPAN? This is given in the problem. This is 3%

What is the probability that exactly 4 people watch CSPAN We are looking for $P(X = 4)$ i.e. four of our neighbors watch CSPAN. X is distributed binomial with $n = 25$ and $p = 0.03$ which can also be written $X \sim Binomial(n = 25, p = 0.03)$. We can use the formula card to see that $P(X = 4) = \binom{25}{4} 4^{0.03} 21^{0.97} = 0.00540$

What is the probabilty that at least one neighbor watches CSPAN? We are looking for $P(X \geq 1) = P(X = 1) + P(X = 2) + \cdots + P(X = 25)$. Probabilities sum to one, so we can rewrite this as: $P(X \geq 1) = 1 - P(X = 0)$. Lets find the probability that $P(X = 0)$. Using the Binomial formula:

$$P(X \geq 1) = 1 - P(X = 0)$$
$$P(X \geq 1) = 1 - \binom{25}{0} 0^{0.03} 25^{0.97}$$
$$P(X \geq 1) = 1 - 0.4669747$$
$$P(X \geq 1) = 0.5330253$$

What is the expected number of neighbors that watch CSPAN? The "expected value" of a binomial is another way of asking for the mean. The mean of the binomial is also given on the formula card: $E[X] = np = 25 \times 0.03 = 0.75$ people on average.

One Proportion Test Assumptions Let's go through a One Proportion problem in all of the ways that it can play out:

> Imagine that your friend says that only 70% of people in Ann Arbor love dogs. You think this is too low so you decide to run a test

Statement: $H_0: p = 0.7, H_a: p > 0.7$ where p is the population proportion of UofM students who love dogs. Lets look at two different scenerios:

- Imagine that you poll 100 people and 90 of them say they love dogs. We first check assumptions:

State	Check	Values	≥ 10
$np \geq 10$	$np_0 \geq 10$	100×0.7	True
$n(1-p) \geq 10$	$n(1-p_0) \geq 10$	$100 \times (1-0.7)$	True

  and since our assumptions are satisfied, we can proceed with the Large Sample Z Test:

$$X \sim Binomial(n = 100, p = 0.7) \text{ (in reality)}$$
$$X \approx N(\mu = 0.7, \sigma = \sqrt{1/100(0.7)(1-0.7)}) \text{ (Approximation)}$$
$$Z \sim N(0,1) \text{ (Distribution of Test Statistic)}$$
$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.9 - 0.7}{\sqrt{\frac{0.7(1-0.7)}{100}}} = 95.23$$

  $P(Z > 95.23) =$ by our z-table ≈ 0

- Now, imagine instead that you only decided to poll 20 people and 18 say they love dogs:

State	Check	Values	≥ 10
$np \geq 10$	$np_0 \geq 10$	20×0.7	True
$n(1-p) \geq 10$	$n(1-p_0) \geq 10$	$20 \times (1-0.7)$	False

  Our assumption is not satisfied. So we need to use the binomial distribution to compute the p-value.

$$X \sim Binomial(n = 20, p = 0.7)$$
$$P(X \geq 18) = P(X = 18) + P(X = 19) + P(X = 20)$$

  which by the binomial formula is:

$$P(X \geq 18) = \binom{20}{18} 18^{0.7} 2^{0.3} + \binom{20}{19} 190.71^{0.3} + \binom{20}{20} 20^{0.7} 0^{0.3}$$
$$P(X \geq 18) = 0.03548313$$

  and this will now be our p-value.

Version: 9000.01
Please email typos to `mtkurzej@umich.edu`