

Review for Exam II

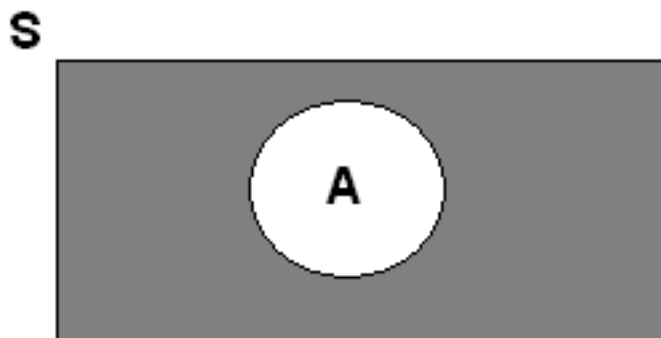
- **Outcome** - a possible result of an experiment
- **Sample Space** - set (collection) of all possible outcomes of an experiment
- **Event** - specific collection of outcomes (results of an experiment that meet certain criteria)

EXAMPLE: An experiment could consist of rolling two dice. An outcome would be the number of dots on each individual die. So if a six and two appeared on the dice, we could consider the outcome to be $(6, 2)$. The sample space would then be all doubles (m, n) , where m and n are some whole number 1 through 6. Possible events could be $E = \{\text{both the numbers are even}\}$, $O = \{\text{both the numbers are odd}\}$, $A = \{\text{both the numbers are less than or equal to 3}\}$, $B = \{\text{both the numbers are greater than 2}\}$.

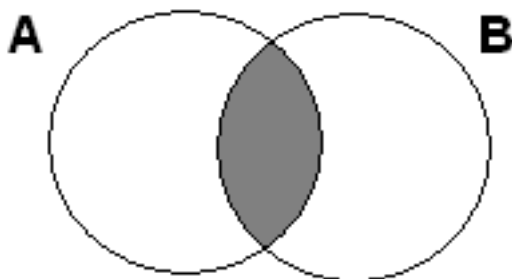
- **Properties of Probability Functions** - probability functions have the following properties:

- (1) the probability of any event is between 1 and 0
- (2) the sum of the probabilities of all the outcomes in a sample space is 1

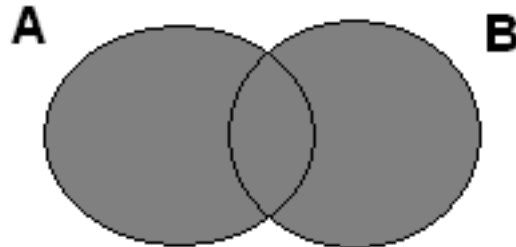
- **Complement** - the complement of an event A , denoted A^c , is everything that is not in event A . This means $P(A^c) = 1 - P(A)$. A picture with A^c shaded is below:



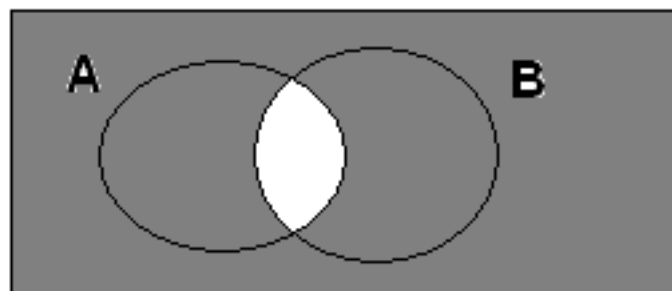
- **Intersection of Events**, \cap - the intersection of events A and B is the same as the event "A and B", that is, it consists of any outcome that is both in event A and in event B . A picture with $A \cap B$ shaded is below:



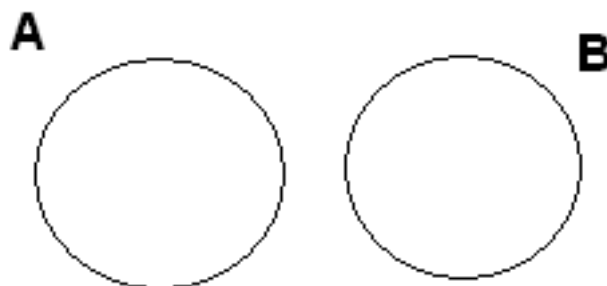
- **Union of Events**, \cup - the union of events A and B is the same as the event "A or B", that is, it consists of any outcome that is in either event A or in event B . A picture with $A \cup B$ shaded is below:



- We can have more complicated events by combination of the above. Below is a picture of $(A \cap B)^c$:



- **Mutually Exclusive** - events A and B are mutually exclusive, or disjoint, if there is no intersection, in other words, if they have nothing in common. So $P(A \cap B) = 0$. A picture of disjoint events A and B is below:



EXAMPLE: Let the events O, E, A, B be defined as above, but for one die instead of two. The events O and E are mutually exclusive. $O \cap B = \{4, 6\}$ while $O \cup B = \{2, 3, 4, 5, 6\}$. $E \cup A = \{1, 2, 3, 4, 6\}$ and $E \cap A = \{2\}$.

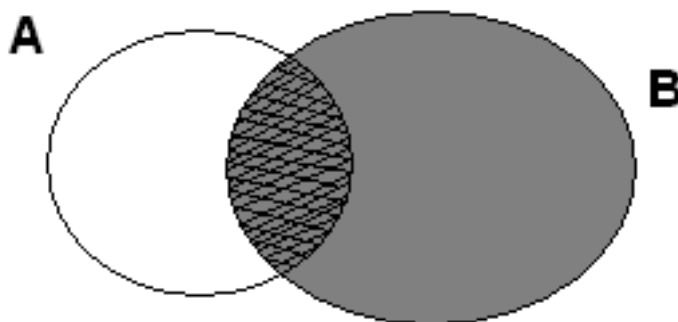
• **Equally Likely** - recall that if events are equally likely and we are there are only a finite number of outcomes, that is $n(\mathbf{S})$ (the number of things in the sample space \mathbf{S}) is some number, we can compute the probability of an event A as

$$P(A) = \frac{n(A)}{n(\mathbf{S})} = \frac{\text{the number of outcomes in } A}{\text{the number of outcomes in } S}$$

• **Conditional Probability** - the conditional probability of A given B , denoted $P(A|B)$, is the probability that A will occur given that the event B happened (so $P(B) \neq 0$). This is the case where we want to think of the "new" sample space as B (since we know B occurred, we restrict ourselves to this case). So since the only part of A that is of interest is the part in B , we have the formula

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

below is a picture of this:



• **Independent Events** - events A and B are independent if one event does not affect the other. Mathematically we can check the independence of events A and B by showing that $P(A \cap B) = P(A) \times P(B)$. Another possibility would be showing that $P(A|B) = P(A)$, that is, B does not affect the probability of A .

• **Probability Rules** - we have the following rules for computing probabilities:

If $A = \{O_1, O_2, \dots, O_k\}$ (so A consists of outcomes O_1 through O_k where these outcomes are all mutually exclusive) then $P(A) = P(O_1) + P(O_2) + \dots + P(O_k)$.

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ if A and B are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$

$P(A \cap B) = P(A|B) \times P(B) = P(B|A) \times P(A)$ if A and B are independent, then $P(A \cap B) = P(A) \times P(B)$. The probability $P(A \cap B)$ is sometimes called the joint probability.

EXAMPLE: (with a table) - We are given the following table of joint probabilities:

	A	B	C
D	0.1	0.3	0.2
E	0.2	0.1	0.1

Since this is a table of joint probabilities, we simply read the values to find the probability of the intersection of events. For example, $P(B \cap D) = 0.3$. To compute the marginal probability of A , we simply add everything in the A column. So $P(A) = 0.1 + 0.2 = 0.3$. Similarly, $P(E) = 0.2 + 0.1 + 0.1 = 0.4$. To compute any conditional probabilities, we simply find the probability of the intersection and the probability of the conditional event.

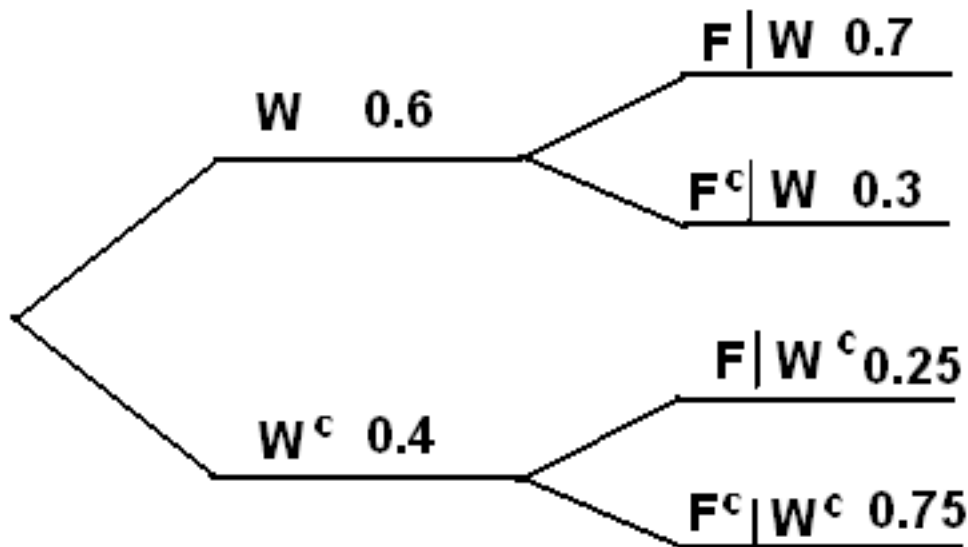
$$P(C|D) = \frac{P(C \cap D)}{P(D)} = \frac{0.2}{0.1 + 0.3 + 0.2} = \frac{0.2}{0.6} \approx 0.33$$

EXAMPLE: (with a tree) Your favorite team is in the final playoffs. You have assigned a probability of 60% that they will win the championship. Past records indicate that when teams win the championship, they win the first game of the series 70% of the time. When they lose the series, they win the first game 25% of the time. The first game is over and your team has lost. What is the probability that they will win the series?

First we will define events. Let W be the event that your team wins the championship and F be the event that they win the first game (so W^c is the event they lose the championship and F^c is the event that they lose the first game). We are given:

$$P(W) = 0.6, P(F|W) = 0.7, P(F|W^c) = 0.25$$

So, using the facts that $P(A) + P(A^c) = 1$ and $P(B^c|A) + P(B|A) = 1$, we can complete the following probability tree:



Now, since we want to compute $P(W|F^c) = \frac{P(W \cap F^c)}{P(F^c)}$, we only need to find $P(W \cap F^c)$ and $P(F^c)$.

$$P(W \cap F^c) = P(F^c|W) \times P(W) = (0.3) \times (0.6) = 0.18$$

We note that $P(F^c) = P(W \cap F^c) + P(W^c \cap F^c)$ (the team either loses the first game and wins the series or loses the first game and loses the series, and these are mutually

exclusive events). We already know $P(W \cap F^c)$, so we only need to find $P(W^c \cap F^c)$.

$$P(W^c \cap F^c) = P(F^c|W^c) \times P(W^c) = (0.75) \times (0.4) = 0.3$$

So $P(F^c) = 0.18 + 0.3 = 0.48$. Thus,

$$P(W|F^c) = \frac{P(W \cap F^c)}{P(F^c)} = \frac{0.18}{0.48} = 0.375.$$

• **Expected Value** - the expected value is given by the formula $\sum x \cdot p(x)$. So if we had the probability table:

x	-3	2	5
$p(x)$	0.3	0.5	0.2

the expected value would be $(-3)(0.3) + (2)(0.5) + (5)(0.2) = -0.9 + 1 + 1 = 1.1$.

• Remember that probability corresponds to area.

• **Binomial Distribution** - the binomial distribution results from an experiment consisting of n trials where each trial results in a success with probability p , or a failure with probability $(1 - p)$ (usually we use 1 for success and 0 for failure). The mean (or expected value) is np , and the variance is $np(1 - p)$ (so the standard deviation is $\sqrt{np(1 - p)}$).

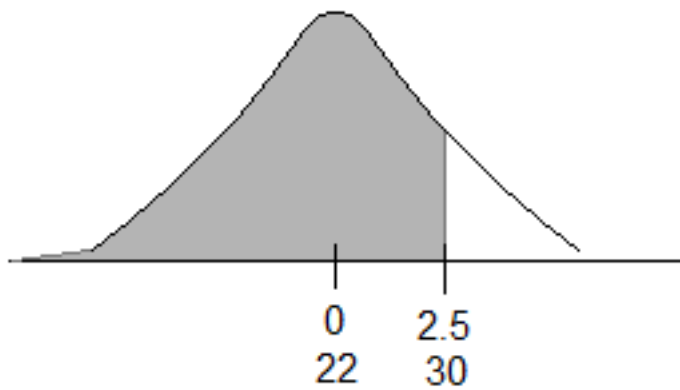
• **Normal Distribution** - this is the distribution with the bell-shaped curve. The normal distributions depends on the mean (usually denoted μ) and the standard deviation (usually denoted σ). We have the table for probabilities of a normal with $\mu = 0$ and $\sigma = 1$ in the back of the book. So, to use this table for *any* normal distribution, we must compute the z -score.

• **Z-score** - the z -score for a particular value of our random variable X is given by the formula $z = \frac{x - \mu}{\sigma}$. Whenever using the table and z -scores to compute probabilities

DRAW A PICTURE so you can see how the probability from the table corresponds to what you are calculating. Most of the time you will have to use the fact that the area on one side of the mean is 0.5, so you will subtract a value from 0.5 or add 0.5 to your value.

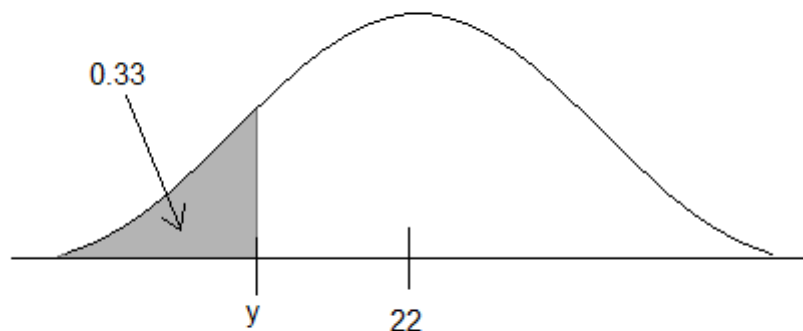
EXAMPLE: The weight of whatsits are normally distributed with a mean of 22 stones and a standard deviation of 4 stones. What is the probability that a whatsit weighs less than 32 stones?

Let X be the weight of the whatsit. The z -score that corresponds to $x = 32$ is $\frac{32-22}{4} = 2.5$. So $P(X < 32) = P(Z < 2.5)$. The picture looks like this:



So we see that we need to add 0.5 to the value from our table. The table gives 0.4938 as the area between 0 and 2.5 under the standard normal curve. So $P(X < 32) = P(Z < 2.5) = 0.5 + 0.4938 = 0.9938$.

EXAMPLE: You should be able to work backwards. For instance, using the previous example. What is the whatsit weight for the 33rd percentile? This means we want to find a value, say y , such that $P(X < y) = 0.33$. The picture would look like this:



So we need to find $0.5 - 0.33 = 0.17$ in our table (the table only gives the area between 0 and a z -score). We look this up in the table to get a z -value of 0.44. However, the

actual value is -0.44 , since we are to the left of the mean, 0 . We now use the formula $x = z\sigma + \mu$ (or solve the z formula for x) to find the value of y .

$$y = (-0.44)(4) + 22 = -1.76 + 22 = 20.24$$

So the 33rd percentile whatsit weight is 20.24 stones.

• **Population Distribution vs. Sample Distribution vs. Sampling Distribution** - for a nice summary of this see slide 8 of Lecture 18.

• **Sampling Distribution of the Sample Mean** - this is also referred to (at least by me in the labs) as the distribution of \bar{X} . When we want to compute the z -score for the mean of a sample of size n we use the formula $z = \frac{x - \mu}{\sigma/\sqrt{n}}$. This comes from ...

• **Central Limit Theorem** - the distribution of the sample mean, \bar{X} , of a random sample of size n from a population with mean μ and standard deviation σ is approximately normal for large n with mean μ (the same as the original population) and standard deviation σ/\sqrt{n} .

EXAMPLE: Back to our whatsit example. What is the probability that a sample of size 9 will have a mean weight greater than 24 stones?

24 corresponds to a z -score of $\frac{24-22}{4/\sqrt{9}} = 1.5$. Using our table (and a picture that YOU made) we can see that our probability is $P(\bar{X} > 24) = P(Z > 1.5) = 0.5 - 0.4332 = 0.0668$.

• **Point Estimators** - for this class point estimators will basically be the sample equivalent. So the point estimator of the population mean, μ , will be the sample mean, \bar{x} , the point estimator of the population variance, σ^2 , will be the sample variance, s^2 , and so on. For a quick discussion about estimators (unbiased, consistent, efficient, etc.) see slides 6 - 11 of Lecture 19.

• **Confidence Interval** - a confidence interval is an interval with an associated level of confidence (some of these definitions are tricky). The level of confidence refers to the probability of the method containing the true population parameter (for us either the mean, μ , or the proportion, p). Once an interval is calculated there is no probability, the true parameter is either in the interval or it isn't. The formulas for the confidence intervals of μ and p are:

$$\bar{x} \pm z \frac{\sigma}{\sqrt{n}} \text{ or } \bar{x} \pm z \frac{s}{\sqrt{n}} \quad \hat{p} \pm z \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Realize how changes affect the interval width, such as increasing n or decreasing the level of confidence decreases the interval width.

EXAMPLE: Compute the 85% confidence interval for the population mean when a sample of 121 observations had a mean of 46 and a standard deviation of 5.2.

Here $\bar{x} = 46$, $s = 5.2$, and $n = 121$. So we still need to find z . Remember, z corresponds to the level of confidence. We want a z such that the area from $-z$ to z is 0.85. So on our table we need to look for $\frac{0.85}{2} = 0.425$. We look for 0.425 in our table to find that $z = 1.44$. So our confidence interval is

$$46 \pm 1.44 \frac{5.2}{\sqrt{121}} \approx 46 \pm 0.681.$$

So we can say that we are 85% confident that the true mean is in the interval (45.319, 46.681).

EXAMPLE: Suppose a department store wants to estimate the average age of its customers, correct to within 2 years, with a 95% level of confidence. The standard deviation is believed to be 8 years. What sample size should they take?

Here we must use one of the formulas for the sample size. If we let B be the allowable range, then the formulas are

$$n = \sigma^2 \left(\frac{z}{B}\right)^2 \text{ for the mean and } n = \hat{p}(1 - \hat{p}) \left(\frac{z}{B}\right)^2 \text{ for the proportion.}$$

Here we are looking for the mean, so we use the first formula. Using the table we find the z that corresponds to an area of $\frac{0.95}{2} = 0.475$ is 1.96.

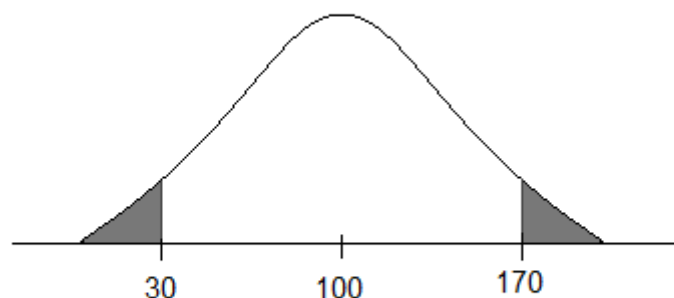
$$8^2 \left(\frac{1.96}{2}\right)^2 = 36(0.9604) = 34.5744$$

So, since we always round up when computing sample sizes, we need a sample of 35 customers.

• **Hypothesis Test** - there are two hypotheses in a hypothesis test, the null hypothesis, H_0 , and the alternative hypothesis, H_1 . Remember, the null hypothesis cannot be proven to be true, we can only show that the alternative hypothesis is true (think of "innocent until proven guilty" where innocence is the null hypothesis).

• **P -value** - the P -value is the probability of seeing something at least as contradictory to the null hypothesis. So we generally want something small (this means that is unlikely that we got a value contradictory to the null hypothesis by chance).

EXAMPLE: Perhaps we are considering the mean of a population. So we will observe the sample mean. Suppose our null hypothesis is $H_0 : \mu = 100$ and our alternative hypothesis is $H_1 : \mu \neq 100$. Suppose we take a random sample and obtain a sample mean of 30. The picture of the P -value is below:



The total area of the shaded regions gives the P -value.