Lecture 12.3: Probability Theory

Chapter 2.4-2.5

One Last Time

You have a new hypothesis testing machine. For any hypothesis test it either concludes:

- "reject H_0 in favor of H_A". Call this a $\oplus$'ve result
- "do not reject H_0". Call this a $\ominus$'ve result.

One Last Time

You have a new hypothesis testing machine. For any hypothesis test it either concludes:

- "reject H_0 in favor of H_A". Call this a $\oplus$'ve result
- "do not reject H_0". Call this a $\ominus$'ve result.

The machine has the following performance specifications:

- $\alpha = \Pr(\text{Reject } H_0 \text{ when } H_0 \text{ true}) = \Pr(\oplus | H_0)$
- Power $= 1 - \beta = \Pr(\text{Reject } H_0 \text{ when } H_A \text{ true}) = \Pr(\oplus | H_A)$

How Reliable Are Your Test Results?

		Test conclusion	
		$\ominus$	$\oplus$
Truth	H_0 true	True Negative (TN)	False Positive (FP)
	H_A true	False Negative (FN)	True Positive (TP)

How Reliable Are Your Test Results?

		Test conclusion	
		$\ominus$	$\oplus$
Truth	H_0 true	True Negative (TN)	False Positive (FP)
	H_A true	False Negative (FN)	True Positive (TP)

- Of the times that we get $\oplus$, the proportion it was right is $\Pr(H_A|\oplus) = \frac{TP}{TP+FP} =$ positive predictive value

How Reliable Are Your Test Results?

		Test conclusion	
		$\ominus$	$\oplus$
Truth	H_0 true	True Negative (TN)	False Positive (FP)
	H_A true	False Negative (FN)	True Positive (TP)

- Of the times that we get $\oplus$, the proportion it was right is $\Pr(H_A|\oplus) = \frac{TP}{TP+FP}$ = positive predictive value
- Of the times that we get $\ominus$ the proportion it was right is $\Pr(H_0|\ominus)$? $\frac{TN}{TN+FN}$ = negative predictive value.

Related Alternative Measure of Reliability

		Test conclusion	
		$\ominus$	$\oplus$
Truth	H_0 true	True Negative (TN)	False Positive (FP)
	H_A true	False Negative (FN)	True Positive (TP)

Related Alternative Measure of Reliability

		Test conclusion	
		$\ominus$	$\oplus$
Truth	H_0 true	True Negative (TN)	False Positive (FP)
	H_A true	False Negative (FN)	True Positive (TP)

The false discovery rate FDR is

$$\text{FDR} = \frac{FP}{TP + FP} = \Pr(H_0 | \oplus)$$

Related Alternative Measure of Reliability

		Test conclusion	
		$\ominus$	$\oplus$
Truth	H_0 true	True Negative (TN)	False Positive (FP)
	H_A true	False Negative (FN)	True Positive (TP)

The false discovery rate FDR is

$$\begin{aligned} \text{FDR} &= \frac{FP}{TP + FP} = \Pr(H_0 | \oplus) \\ &= 1 - \frac{TP}{TP + FP} = 1 - \text{Positive Predictive Value} \end{aligned}$$

Multiple Testing

Recall we use Bonferroni Correction to account for multiple testing (think jelly beans and acne): if we are doing k tests, we use

$$\alpha^* = \frac{\alpha}{k}$$

in our hypothesis tests.

Multiple Testing

Recall we use Bonferroni Correction to account for multiple testing (think jelly beans and acne): if we are doing k tests, we use

$$\alpha^* = \frac{\alpha}{k}$$

in our hypothesis tests. But when you have a lot of tests, like genetic studies, k is very large so α^* gets too small.

Multiple Testing

Recall we use Bonferroni Correction to account for multiple testing (think jelly beans and acne): if we are doing k tests, we use

$$\alpha^* = \frac{\alpha}{k}$$

in our hypothesis tests. But when you have a lot of tests, like genetic studies, k is very large so α^* gets too small. You will almost never reject H_0 and hence get a lot of false negatives.

Multiple Testing

Recall we use Bonferroni Correction to account for multiple testing (think jelly beans and acne): if we are doing k tests, we use

$$\alpha^* = \frac{\alpha}{k}$$

in our hypothesis tests. But when you have a lot of tests, like genetic studies, k is very large so α^* gets too small. You will almost never reject H_0 and hence get a lot of false negatives.

As an alternative, you can try to control the FDR.

Controlling the False Discovery Rate

Example: let's say you're using microarrays to compare expression levels for 100,000 genes between liver tumors and normal liver cells.

Controlling the False Discovery Rate

Example: let's say you're using microarrays to compare expression levels for 100,000 genes between liver tumors and normal liver cells.

You're going to do additional experiments on any genes that show a significant difference between the normal and tumor cells, and you're willing to accept up to 10% of the genes with significant results being false positives.

Controlling the False Discovery Rate

Example: let's say you're using microarrays to compare expression levels for 100,000 genes between liver tumors and normal liver cells.

You're going to do additional experiments on any genes that show a significant difference between the normal and tumor cells, and you're willing to accept up to 10% of the genes with significant results being false positives.

i.e. You'll find out they're false positives when you do the followup experiments.

Controlling the False Discovery Rate

Example: let's say you're using microarrays to compare expression levels for 100,000 genes between liver tumors and normal liver cells.

You're going to do additional experiments on any genes that show a significant difference between the normal and tumor cells, and you're willing to accept up to 10% of the genes with significant results being false positives.

i.e. You'll find out they're false positives when you do the followup experiments.

In this case, you would set your false discovery rate to 10 percent.

Random Variable

A random process or variable with a numerical outcome is called a random variable, and is typically denoted by an upper case letter. E.g. X, Y, or Z

Intuitively Thinking: Expected Value

Easy example: Coin flips. Say we flip a fair coin $n = 10$ times with probability $p = \frac{1}{2}$ of heads.

Intuitively Thinking: Expected Value

Easy example: Coin flips. Say we flip a fair coin $n = 10$ times with probability $p = \frac{1}{2}$ of heads.

How many heads do you expect to get?

$n \times p = 10 \times \frac{1}{2} = 5$

Intuitively Thinking: Expected Value

Slightly more complicated example: Say you have a random variable X:

x	2	3	4	10	11
$\Pr(X = x)$	$\frac{15}{100}$	$\frac{25}{100}$	$\frac{10}{100}$	$\frac{30}{100}$	$\frac{20}{100}$

E.g. We observe $X = 3$ with prob .25

Intuitively Thinking: Expected Value

Slightly more complicated example: Say you have a random variable X:

x	2	3	4	10	11
$\Pr(X = x)$	$\frac{15}{100}$	$\frac{25}{100}$	$\frac{10}{100}$	$\frac{30}{100}$	$\frac{20}{100}$

E.g. We observe $X = 3$ with prob .25

Is the value we expect to observe:

$$\frac{2+3+4+10+11}{5} \quad = \quad 6\ ?$$

Intuitively Thinking: Expected Value

No, each of the x's have different probability of occurring.

Intuitively Thinking: Expected Value

No, each of the x's have different probability of occurring.

For each x, we assign weight $\Pr(X = x)$.
i.e. for all x, we have $x \cdot \Pr(X = x)$:

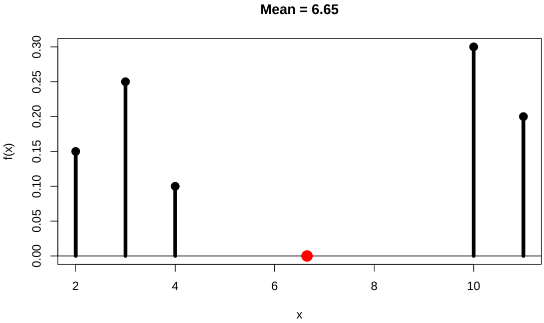
Expected Value

The expected value is a weighted average of all possible values x. This can be thought of as a measure of center:

This is also called the mean and expectation of X. Typically denoted by μ.

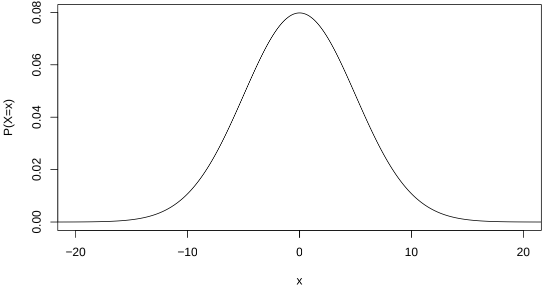
Expected Value

You can also think of the mean as the center of mass or balance point (marked with red point):



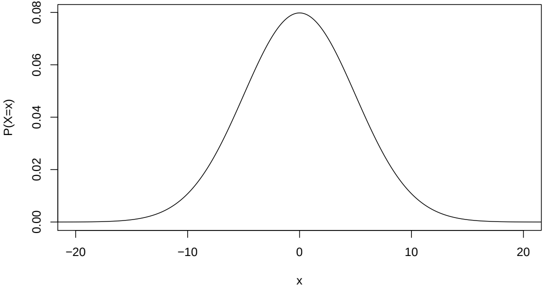
Intuitively Thinking: Measures of Spread

Consider the following (continuous) distribution with $\mu = 0$. Let's build a measure of expected "spread".



Intuitively Thinking: Measures of Spread

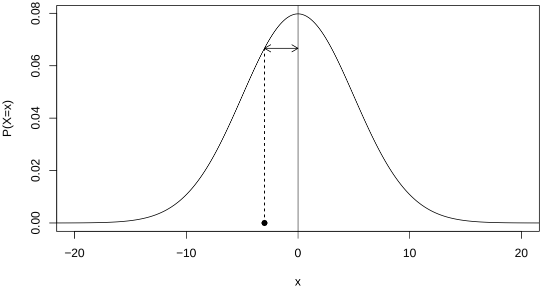
Consider the following (continuous) distribution with $\mu = 0$. Let's build a measure of expected "spread".



Let's define "spread" as the absolute deviation from μ: $|x - \mu|$.
i.e. $+$'ve & $-$'ve deviations of the same magnitude are treated the same.

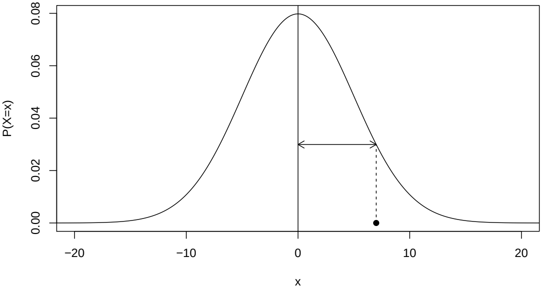
Intuitively Thinking: Measures of Spread

When $x = -3.0$, the abs. deviation from μ is $|-3.0 - \mu| = 3.0$.
Note $P(X = x) = 0.066$.



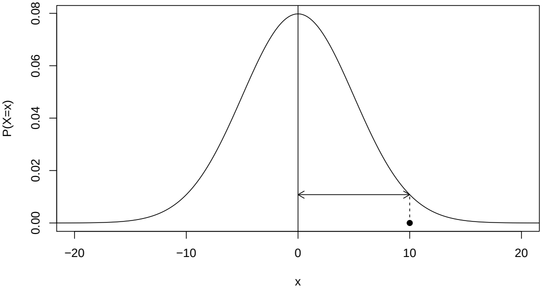
Intuitively Thinking: Measures of Spread

When $x = 7.0$, the abs. deviation from μ is $|7.0 - \mu| = 7.0$.
Note $P(X = x) = 0.030$.



Intuitively Thinking: Measures of Spread

When $x = 10.0$, the abs. deviation from μ is $|10.0 - \mu| = 10.0$.
Note $P(X = x) = 0.011$.



Intuitively Thinking: Measures of Spread

So say we do this for all x and take a weighted average of the $|x - \mu|$ where the weights are $P(X = x)$.

Voilà: Our notion of expected spread.

Variance

The variance σ^2 AKA Var(X) of a distribution is

Variance

The variance σ^2 AKA Var(X) of a distribution is

It is the expected squared deviation from the mean, and not absolute deviation (like in our example). i.e. not

Variance

The variance σ^2 AKA Var(X) of a distribution is

It is the expected squared deviation from the mean, and not absolute deviation (like in our example). i.e. not

Why square? Treats $+$'ve and $-$'ve deviations as the same, but also easier to do calculus on x^2 than $|x|$.

Estimators

In statistics, an estimator is a rule for calculating an estimate of a given quantity based on observed data: thus the rule and its result (the estimate) are distinguished.

Estimators

In statistics, an estimator is a rule for calculating an estimate of a given quantity based on observed data: thus the rule and its result (the estimate) are distinguished.

Recall how we've said $\overline{x}$ is a point estimate of μ. $\overline{x}$ is based on observed data.

Estimators

In statistics, an estimator is a rule for calculating an estimate of a given quantity based on observed data: thus the rule and its result (the estimate) are distinguished.

Recall how we've said $\overline{x}$ is a point estimate of μ. $\overline{x}$ is based on observed data.

Before we've observed the data, $\overline{X}$ is still random and is the estimator of μ.

Estimators

In statistics, an estimator is a rule for calculating an estimate of a given quantity based on observed data: thus the rule and its result (the estimate) are distinguished.

Recall how we've said $\overline{x}$ is a point estimate of μ. $\overline{x}$ is based on observed data.

Before we've observed the data, $\overline{X}$ is still random and is the estimator of μ.

In general, we have an unknown population parameter θ and an estimator $\widehat{\theta}$.

Sample Mean as an Estimator

Bias

One property we want our estimators to have is unbiasedness. i.e.

i.e. we expected the estimator's value to be the unknown parameter.

Recall from Lecture 1.3

One example of a non-representative sample is a biased sample. For example, convenience samples are samples where individuals who are easily accessible are more likely to be included.

Recall from Lecture 1.3

1. The Royal Air Force wants to study how resistant their airplanes are to bullets. They study the bullet holes on all the airplanes on the tarmac after an air battle against the Luftwaffe (German Air Force).
2. I want to know the average income of Reed graduates in the last 10 years. So I get the records of 10 randomly chosen Reedies. They all answer and I take the average.
3. Imagine it's 1993 i.e. almost all households have landlines. You want to know the average number of people in each household in Portland. You randomly pick out 500 phone numbers from the phone book and conduct a phone survey.
4. You want to know the prevalence of illegal downloading of TV shows among Reed students. You get the emails of 100 randomly chosen Reedies and ask them "How many times did you download a pirated TV show last week?"