Introduction to classical statistics 1

Jonathan Shapiro

Department of Computer Science

Announcements

Why I am on strike (and why I am not on strike today)

1. They are trying to reduce our pensions. 2. Many of my colleagues are on precarious, short-term contracts, which the university will not even acknowledge. 3. Pay inequality ♀/♂ 4. Workload (overload) is growing.

Please email the powers that be

Use the QR-code, or president@manchester.ac.uk and patrick.hackett-REGISTRAR@manchester.ac.uk, saying: 1. I support the strike. Please push for negotiations to end this strike so our lecturers are treated properly. 2. I don't support the strike, and it is robbing me of the education I pay dearly for. Please sort it. 3. Or whatever you feel like.

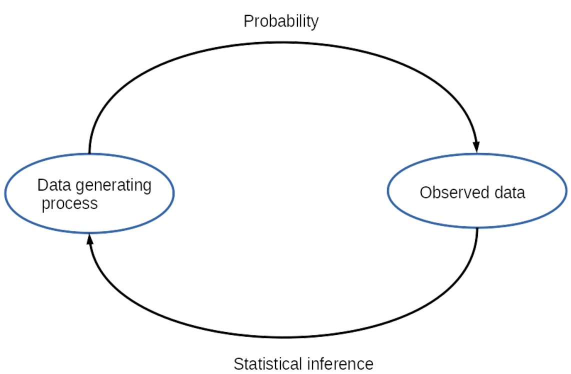
MANCHESTER 1824
The University of Manchester

What is statistics?

"Statistics" used to be called "inverse probability"

Probability theory predicts the observed data given a generating model;

Statistics starts with the data and tries to predict the model which generated it.



Why do I say "classical" statistics?

Classical statistics: Invented by R.A. Fisher, Pearson, Neyman, and others in the 1920's before computers. Uses asymptotic distributions. What most often meant by "statistics". Uses "frequentist" probability.

Bayesian statistics: "Invented" by Bayes, but independently and more thoroughly by Laplace, later by Jeffreys and Jaynes. Covered in later lectures.

Statistics is about hypothesis testing using data

- A fundamental aspect of science is creating hypothesis.
- Testing hypothesis often using data.

What is a hypothesis?

A hypothesis is a predicate (true or false) statement.

```
- To be a scientifically valid, a hypothesis must have other properties.
```

A scientifically valid hypothesis must be testable

It must make predictions or have consequences which can be tested.

```
Negative examples:

1. I am very intelligent, but if you try to measure it, my intelligence stops worki
ng.

2. If a tree falls in the forest, and there is no one nor measuring devices to dete
ct it, it makes no sound. Otherwise, it makes a sound.
```

A scientifically valid hypothesis must be falsifiable

- There must exist an experiment which could have an outcome which shows the hypothesis to be false.
- The necessity of falsifiability introduced by Karl Popper. (Used to criticize Marx, Freud, Adler, and others.)

Example

I can read the mind of any person in this room, but only if there are no skeptics in the room.

Empirical science (data) can never "prove" a hypothesis is true.

It can disprove the hypothesis by getting evidence that contradicts the hypothesis.

1. All coins are "fair" (probability of Heads = probability of tails)

   ```
   Apparently not true for the first Belgian Euro coin.
   ```

2. All species of the birds can fly.

   ```
   Until a flight-less species was encountered, this could have been true.
   ```

A concept for hypothesis testing

- Try every way you can think of to *disprove* your hypothesis
- The more you try and fail, the more evidence you have gained that the hypothesis is true.
- The more evidence the hypothesis explains or predicts, the more likely it is to be true.

Comparing two hypotheses is very effective for hypothesis testing

- The data is more probable assuming hypothesis A than it is when assuming hypothesis B.
- Requires a probability function for the data for each hypothesis.
- Requires a definition of "more" probable. How much more is sufficient?

MANCHESTER 1824
The University of Manchester

Basics of Classical Statistics

A very common scenario where classical statistics can be applied.

Two quantities A and B each derived from two different data sources. E.g. the mean of each dataset.

Possible hypotheses,

1. A and B are the same
2. A < B
3. A is different from B
4. A > B

Examples

1. Is the recovery rate using a new drug A, higher than that of the existing drug B?

1. Is the new algorithm A faster than the existing, state-of-the-art algorithm B?

The standard description

Null hypothesis: (H_0) The hypothesis that there *is* no difference; perceived differences are due to chance alone. (A = B)

Alternative hypothesis: (H_1) The difference is real. At a certain probability level α.

The null hypothesis comes from a known probability function (usually normal). From this you compute whether the likelihood of the null hypothesis is

1. Less than α - reject the null hypothesis, the difference is "statistically significant"

1. Greater than α - "fail to reject the null hypothesis"; the differences are not significantly different.

MANCHESTER 1824
The University of Manchester

A simple scenario: comparing data to a known value

Comparing quantity derived from data to a known value K

1. Is A consistent with K?
2. Is A different from K?
3. Is A < K?
4. Is A > K?

Back

Example: coin-flipping experiment

- We hear of a coin said to have a different probability of heads and tails.

- We investigate this statistically.

Pick an α value

$\alpha = 5\%$ is commonly used.

1. Then you can be 95% sure they are different, if they are.

1. If you want to be more sure they are different, choose a lower value, e.g. $\alpha = 1\%$.

1. Higher values are also possible.

Choose the null hypothesis

What would be the null hypothesis?

The null hypothesis H_0

The coin is a fair coin.

1. The probability of heads is the same as the probability of tails, namely 0.5, or 50%.
2. Any differences are due to randomness.

Choose the alternative hypothesis.

Alternative Hypotheses

The alternative hypothesis H_1

The probability of heads for this coin is $\neq 0.5$.

Get the data

Flip the coin 100; it comes up heads 41 times

Is it a fair coin?

Is it a fair coin?

- Because we chose $\alpha = 0.05$ and H_1 this coin's probability of heads $\neq 0.5$,
- If the probability of 41 heads in 100 flips is
  - less than 0.05, it is not a fair coin
  - else we cannot reject the hypothesis that it *is* a fair coin.

This is called a "two-tail" probability problem.

```
- We don't care if the probability is in the upper tail or the lower tail.

- I.e. we don't care if it is greater than or less than $0.05$.

- You'll see.
```

How do we choose between H_0 and H_1?

Next section

MANCHESTER 1824
The University of Manchester

Confidence intervals and tail probabilities

Confidence interval

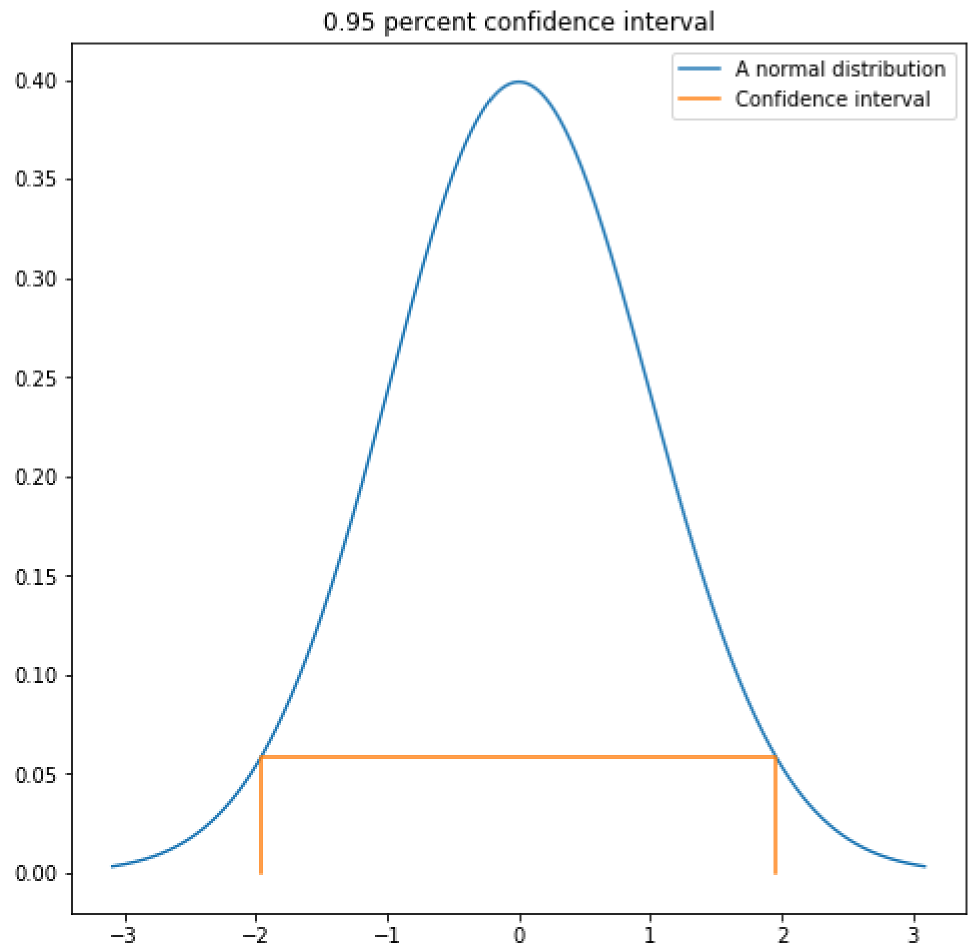
$(1-\alpha)$**-confidence interval:** is the shortest interval which contains $1-\alpha$ percent of the probability mass

In our case, we are interested in $\alpha = 5\%$, we are interested in the 95% confidence interval.

1. If our "test statistic" (41 heads) lies *outside* appropriate confidence interval, reject H_0.

1. If our "test statistic" (41 heads) lies *inside* appropriate confidence interval, we cannot reject H_0.

95%-confidence interval of a Normal distribution

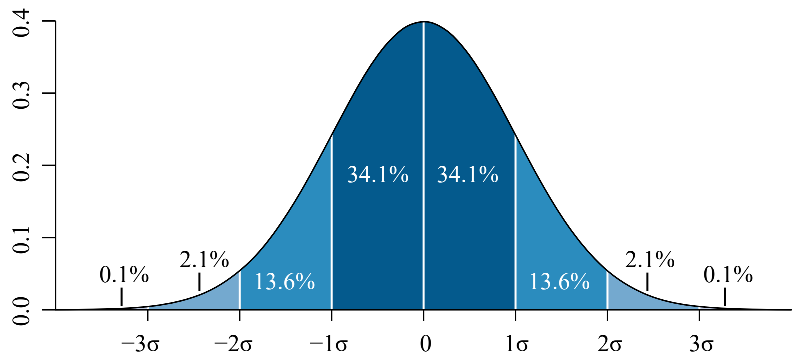


Back

What if our hypothesis was coin heads probability < 0.5?

We need to find the value of the null probability distribution such that the mass of the probability less that value is α.

Tails of the Normal distribution



File:Standard deviation diagram.svg from Wikimedia Commons

We can use the cumulative distribution function (CDF)

Cumulative distribution function CDF:

1. Call it $F(x)$.

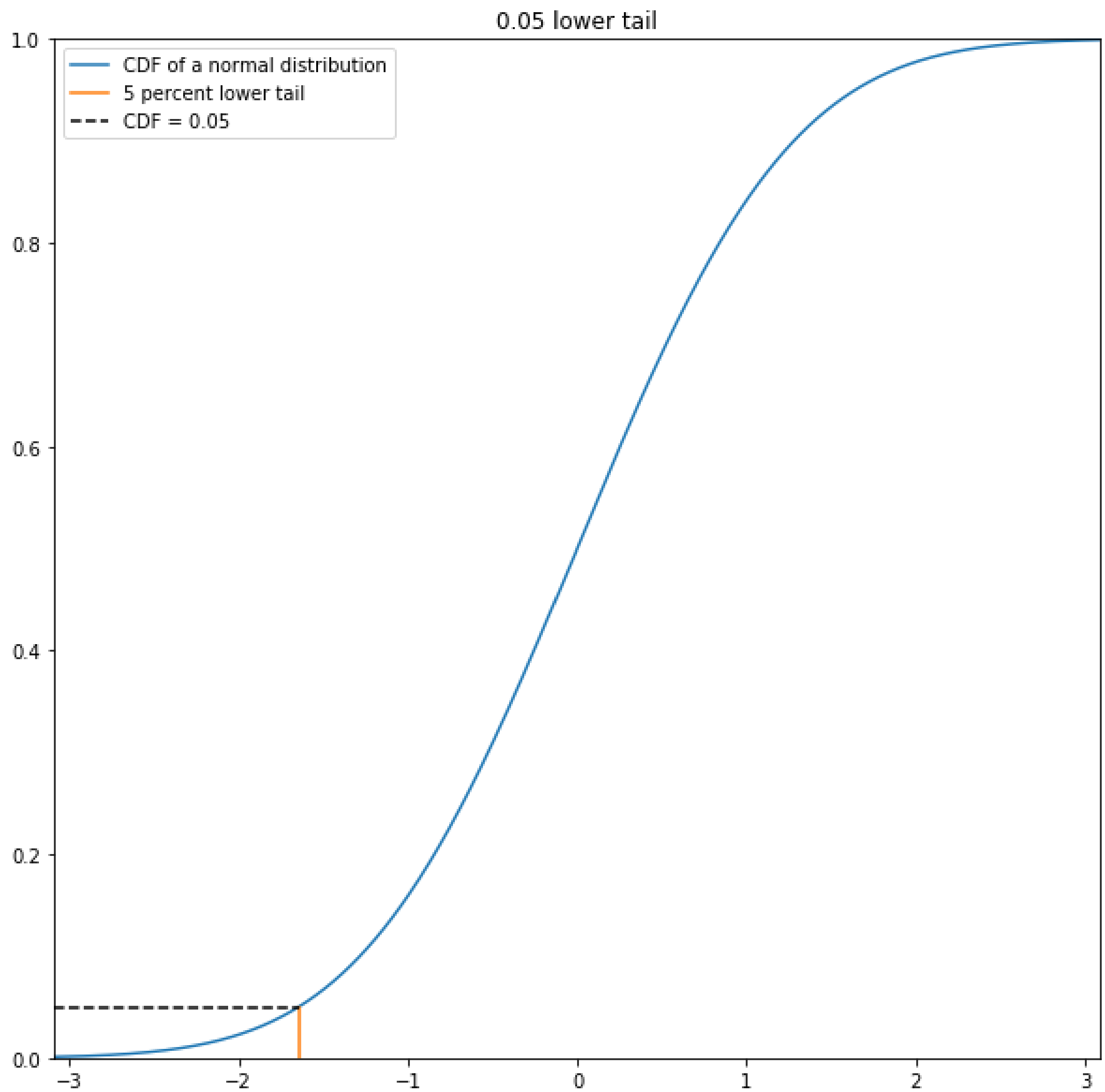
1. Gives the probability that a sample from a probability distribution v $f(v)$ is less than x

1. $$F(x) = \int_{-\infty}^{x} f(v)dv.$$

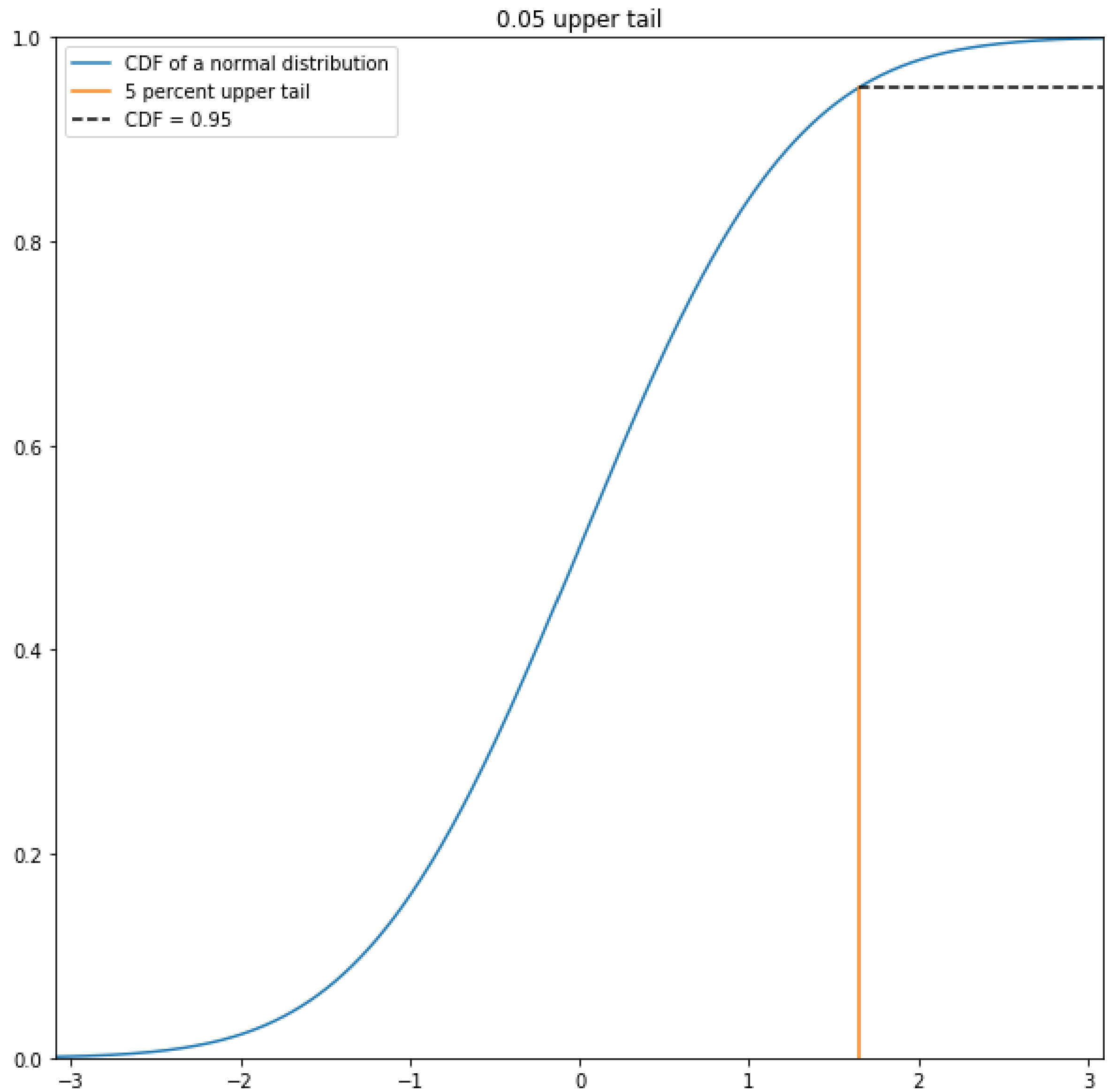
Want to find the critical x_{crit} so that $F(x_{\text{crit}}) = \alpha$.

Inverse cumulative distribution function (PPF)

Called PPF for `percentage point function` in python package `scipy.stats` .



Elsewhere abbreviated ICDF.


0.05 upper tail
CDF of a normal distribution
5 percent upper tail
CDF = 0.95
1.0
0.8
0.6
0.4
0.2
0.0
−3
−2
−1
0
1
2
3

- To get the lower tail, use $\mathrm{PPF}(x) = \alpha$.
- To get the upper tail, use $\mathrm{PPF}(x) = 1 - \alpha$

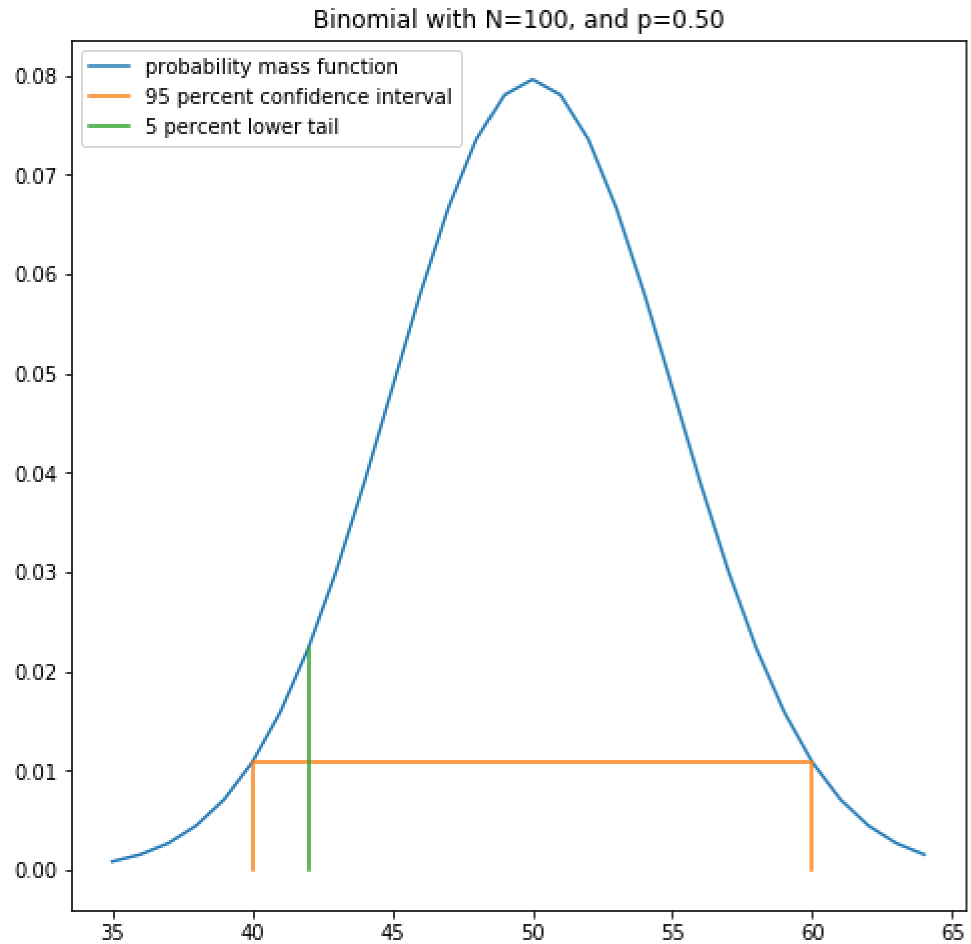
Back to: is it a fair coin?

Is it a fair coin?

For a coin-flipping problem, assuming a fair coin what is the appropriate distribution for the null hypothesis?

It is binomial distribution with $p = 0.5$ and $N = 100$

The result from the binomial probability mass function



[Gaussian]

Summary of the result

If the assumption of H_1 is probability of heads for the coin is different than 0.5, then we cannot rule out the possibility that it is a fair coin at the 5% level.

If the assumption of H_1 is probability of heads for the coin is less than 0.5, then we can. The probability of head is statistically significantly lower than 0.5.

Slightly paradoxical.

MANCHESTER 1824
The University of Manchester

Summing up

After studying this lecture, the student should be able to:

1) Distinguish a good hypothesis from a bad one

2) Given a scenario, produce the null hypothesis and an alternative hypothesis.

3) Interpret the situation when the null hypothesis is not ruled out and when it is.

4) Start to understand how to determine whether the mean or similar of a dataset differs from a known value.

The end