6. Statistics and modelling
Statistical modelling and the central dogma of statistics and probability

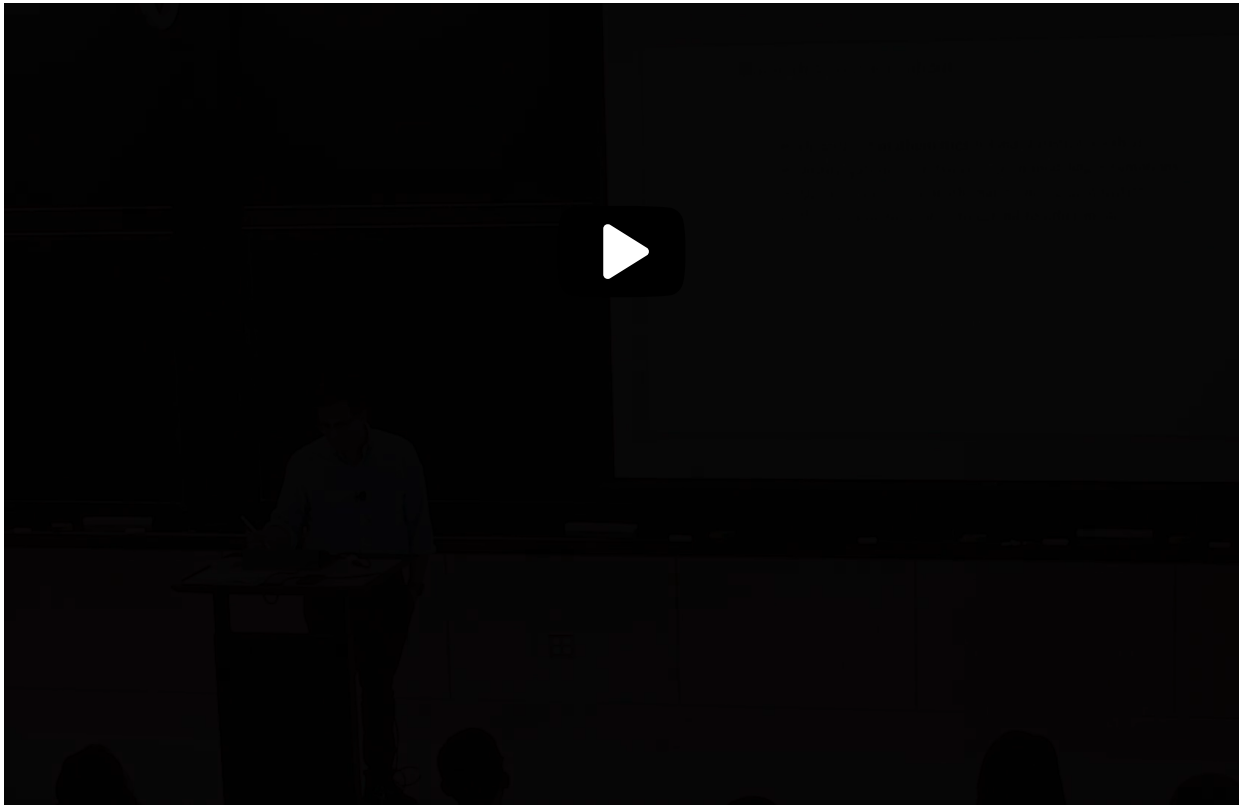
that we have with 95% confidence.

So hopefully, those things mean nothing to you,

and we'll know exactly what we mean by 95% percent confident,

for example.

And what does it mean to think about



And what does it mean to think about other studies also?

That's the frequentist point of view,

and we'll get back to that.

End of transcript. Skip to the start.

9:43 / 9:43 1.50x

Video

Download video file

Transcripts

Download SubRip (.srt) file

Download Text (.txt) file

Probability or statistics 1

2/2 points (graded)

Determine whether each of the problems below is a probabilistic or a statistical problem. (You are **not** asked to solve them.)

1. Assume we have a population consisting of two subpopulations, A and B. A particular drug has a different chance of treatment success depending on the subpopulation, namely 70% for group A and 50% for group B.

   Assume that subgroup A is 10% of the entire population and subgroup B is 90%. What is the chance of a successful treatment if we pick a random person from the entire population?

   This is a

   - statistical problem
   - probabilistic problem ✔

2. Now, consider the scenario where we do not know the true composition of the population, which may be different from the previous setup. Among 1000 randomly chosen patients, we observe that the treatment was successful in 700 of them. What is a good estimate of the composition of the population?

   This is a

   - statistical problem ✔
   - probabilistic problem

Solution:

1. The first one is a probabilistic problem, because we are given all relevant parameters and are trying to compute corresponding derived probabilities. In particular, if we denote the subpopulation of a randomly selected person by $X \in \{A, B\}$ and the treatment outcome of drug D by $Y \in \{\text{sucess}, \text{failure}\}$, we are given $\mathbf{P}(X = A), \mathbf{P}(X = B), \mathbf{P}(Y = \text{success} | X = A)$, and $\mathbf{P}(Y = \text{success} | X = B)$ and are asked to compute $\mathbf{P}(Y = \text{success})$.

2. The second one is a statistical problem, because we are trying to estimate an underlying probabilistic parameter from data. More explicitly, we have 1000 i.i.d. draws from the Bernoulli random variable Y, 700 of which correspond to $Y = \text{success}$, and are now being asked to draw conclusions about $\mathbf{P}(Y = \text{success})$ and from there about $\mathbf{P}(X = A)$ and $\mathbf{P}(X = B)$.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Probability or statistics 2

2/2 points (graded)

John Arbuthnot wrote a paper in 1710 entitled 'An Argument for Divine Providence', where he studied, based on the Christening records in London for 1629-1710, the chances that a randomly chosen baby born is a girl or a boy. Is this a statistical problem, or a probabilistic problem?

- (●) A statistical problem.

A probabilistic problem.

✔

Next, you read Arbuthnot's paper, and went to a gyneacology facility, in which there are 10 babies whom are expected to be born on the day you arrived, and you are interested in, what are the odds that 6 of those will be a boy, and the remaining will be a girl. Is this a statistical problem, or a probabilistic problem?

A statistical problem.

A probabilistic problem.

✔

Solution:

The first one is a statistical problem, and the second one is a probabilistic problem. To see this, suppose that each newborn baby is a boy with probability p, and a girl with probability $1-p$. Suppose also that the sex of each newborn is independent of the sex of all others. The data that Arbuthnot analyzed simply corresponds to realizations of this Bernoulli variable; and from that knowledge, we simply want to extract the underlying parameter, p. This is an example of a statistical problem.
You take Arbuthnot's finding, and assume that this is the 'true' probability for the aforementioned birth process; and want to compute a probability, which is simply

$$\binom{10}{6} p_A^6 (1-p_A)^4,$$

where p_A is the value that Arbuthnot has reported (namely, you are computing the probability that a certain Binomial random variable is equal to 6.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Probability or statistics 3

2/2 points (graded)

A doctor realizes that there is an allergy medicine which is effective in treating seasonal allergies with probability at least 90%. From here, he claims:

- Out of 100 patients admitted to clinic with seasonal allergies, this drug will cure 90 patients, on average.
- At least 70 patients will be cured, with 99.99% chance.

Does he rely on statistics, or probability?

- Statistics
- Probability ✔

Now, a newly-hired scientist at a pharmacology company performs an experiment, and based his observations, deduces that, “I am 95% confident that if we repeat this experiment, then the drug will be effective on between 85% and 95% patients." Does he rely on statistics, or probability?

- (●) Statistics
- ( ) Probability

✔

Solution:

- The doctor relies on probability. The point of discussion is about averages and the odds that at least 70 patients will be cured.
- The scientist relies on statistics, since is using observations This is hinted at, because he discusses confidence regions.

Submit You have used 1 of 2 attempts

ℹ Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 1 Introduction to statistics:Lecture 1: What is statistics / 6. Statistics and modelling

Add a Post

Show all posts ▼ by recent activity ▼

💬	GPA and Central dogma of biology For those (like me) that didn't know it... "GPA stands for Grade Point Average. It is a standard way of measuring academic achievement in the U.S." "T...	2
💬	The solution to Question 2 is not clear. the solution states that "suppose that each newborn baby is a boy with probability p, and a girl with probability 1−p, independent of each other". If w...	6
💬	The last question is not worded properly The statement does not specify if the new hire has observed anything, looks more like he/she is just taking the 0.9 probability as a fact and then calc...	9
💬	Probability vs Statistics Maybe not the most complete definition, but what make me think to the difference... **Probability**: Given this stochastic model (i.e. *assuming* th...	2
💬	Free will and randomness Free will maybe an example of *complicated processes* which are *simpler processes + random noise*. And the reason as we are not deterministic ...	3
?	Probability or Statistics 3, first question I can't seem to understand why this is a probability question or statistics question. Don't you require statistics in order to infer that "At least 70 patie...	4

Learn About Verified Certificates