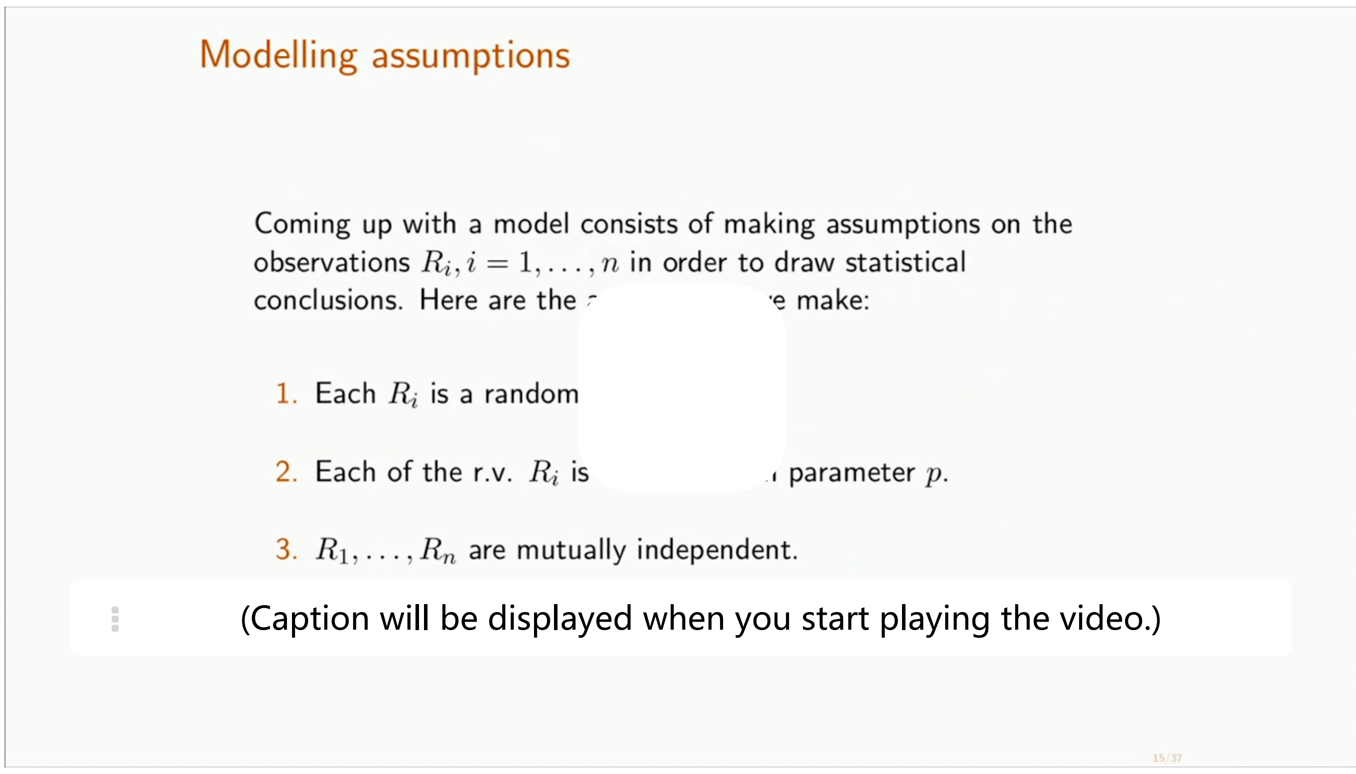
9. The first example: modelling assumptions
The kiss example: modelling assumptions



And the probability that Ri is equal to 0 is 1 minus p.

That's the Bernoulli random variable.

Then the other thing that I'm assuming,

which is something that's always hidden--

and the reason why I'm making this assumption

is because I want to be able to use tools from probability.

And remember the rules for intersection, the probability a

and b is the probability of a times the probability of b

relies on the fact that a and b are independent events.

Remember that one?

Then this is the kind of stuff we're

going to be using all the time.

And if a and b are not independent,

视频
下载视频文件

字幕
下载 SubRip (.srt) file
下载 Text (.txt) file

Independence

1/1 point (graded)
Consider a probabilistic experiment where we roll a dice and toss a coin. We compute the probability that the dice gives 5 and the coin lands Heads: $1/6 \cdot 1/2 = 1/12$. What assumptions are we implicitly using in this calculation? Choose all that apply.

- [x] Each dice roll is uniformly distributed within the set $\{1, 2, 3, 4, 5, 6\}$ and each coin toss is uniformly distributed in $\{\text{Heads}, \text{Tails}\}$. ✔
- [x] The dice roll and coin toss are independent. ✔
- [ ] The random variables corresponding to outputs of each of these experiments are i.i.d.

✔

Solution:

The correct answers are the first and second choices.

Let X denote the output of the dice roll and Y denote the output of the coin toss. We are looking at the probability

$$\begin{aligned} \mathbb{P}(X = 5, Y = \{Heads\}) &= \mathbb{P}(X = 5)\,\mathbb{P}(Y = \{Heads\}) \\ &= \frac{1}{6} \cdot \frac{1}{2}. \end{aligned}$$

The first line, where we express the joint probability as a product, uses the fact that coin toss and dice roll are independent. The second line, where we substitute the values $1/6$ and $1/2$, uses the uniformity assumption to explicitly compute these probabilities.

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

(Optional) Examples of I. I. D. variables

0 points possible (ungraded)
Remember from the course, *Introduction to Probability*, that **i.i.d.** stands for **independent and identically distributed** .

A collection of random variables $X_1, \ldots, X_n$ are **i.i.d.** if each X_i follows a distribution $\mathbf{P}_i$, and all those distributions are the same, and (apart from having the same distribution), each X_i does not contain information about the other realizations.

Decide which of the following collections are (approximately) i.i.d. (independent and identically distributed).
(Choose all that apply.)

- [x] People selected randomly (with replacement) by their address from a directory. ✔
- [ ] The first two consecutive words of a random page in a book.
- [x] Repeated dice rolls of the same die. ✔
- [ ] Temperature measurements on Monday and Tuesday in the same week.

✔

Solution:

If we select people randomly from a base population, we are in charge of the sampling and can do so in an independent manner. Since the distribution is the same, this is a case of i.i.d. random variables. Note that if the population is large, the distribution of a small number of draws actually behaves similar to an i.i.d. draw, even if we sample without replacement.

Words in text documents are not independent because they follow certain compositional rules. For example, it is likely to find a noun preceded by an article.

If a dice is rolled repeatedly, we consider each roll an independent draw from the same distribution, hence this is an iid process.

Temperature measurements are highly correlated in time, although winter in Boston, where MIT is can sometimes make you think otherwise. Roughly speaking, if Monday has a warm weather, you would probably not expect Tuesday to be freezing cold.

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

讨论

显示讨论

主题： Unit 1 Introduction to statistics:Lecture 1: What is statistics / 9. The first example: modelling assumptions