The background of the slide features a stylized globe on the left side, partially obscured by a dense pattern of binary code (0s and 1s) that recedes into the distance, creating a sense of depth and digital connectivity. The colors are primarily light blue and white.

Foundations of Data Science: Hypothesis testing - Principle of hypothesis testing

What is hypothesis testing?

1. Decide whether a hypothesis or model is compatible with data from observational studies or randomised experiments
2. Investigating mechanisms specific to data.

Method of hypothesis testing

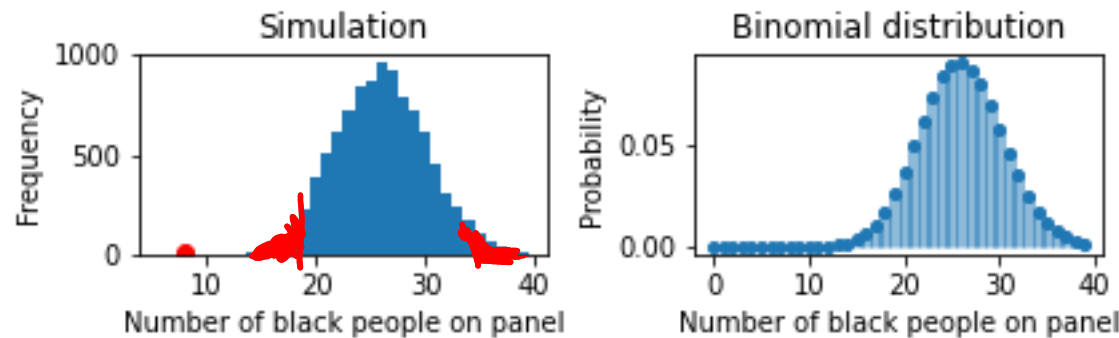
The null hypothesis is H_0 : The claim we initially assume to be true, formulated as a statistical model.

The alternative hypothesis is H_a : The claim that is contradictory to H_0 , typically not formulated as a statistical model.

Aim: Reject or not reject H_0

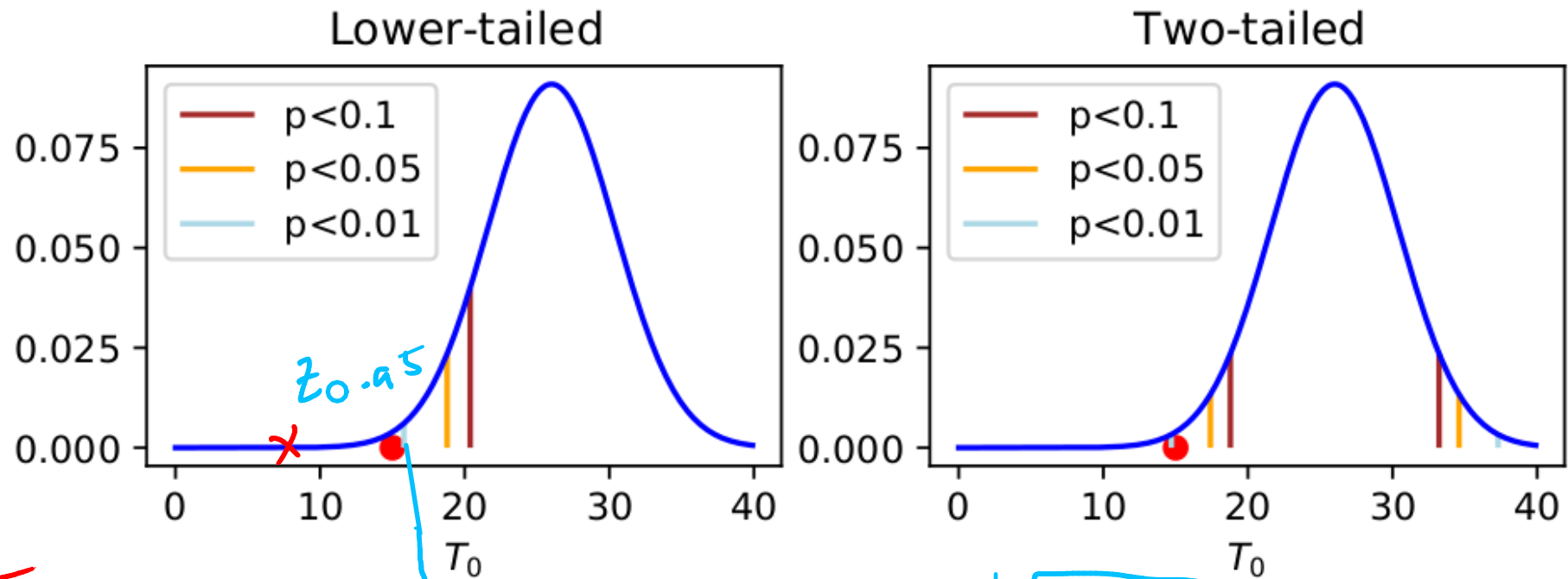
Test procedure

1. Test statistic : e.g number of Black people on jury panel
2. Distribution of test statistic under H_0



3. (a) Rejection region
(b) Return a p-value

Rejection regions



$$T_0 = 15$$

$$\mu + \sigma z_{0.95} = np + \sqrt{np(1-p)} z_{0.95}$$

"We reject H_0 at the 1% level"

n large enough \Rightarrow Binomial dist \approx Normal dist
 $\mu = np$, $\sigma^2 = np(1-p)$
 $\Rightarrow z = \frac{T_0 - \mu}{\sigma}$ is normally distributed