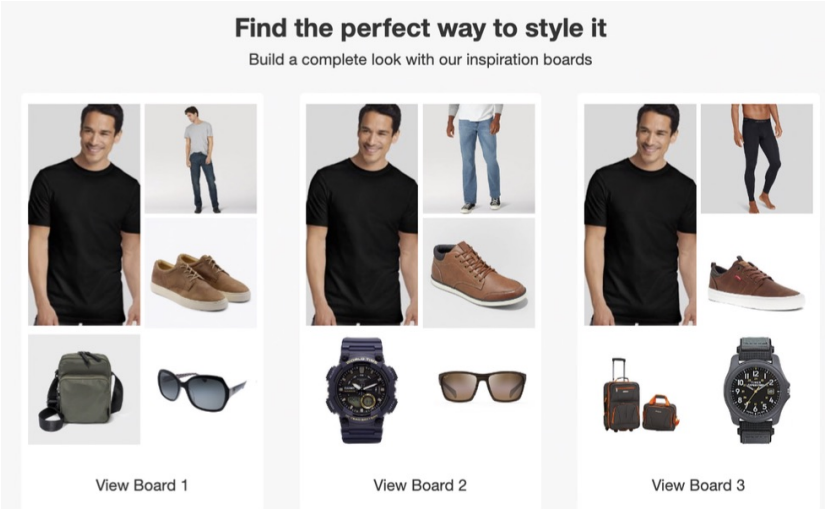
Hypothesis Testing

3

— Interview Question

Recommender System



When a customer buys a T-shirt, a recommender algorithm also suggests a few related items The recommender system in production (legacy) that has a success rate of 10%

You and your team have developed a new deep learning algorithm for recommendation
It is tested before deploying.

Of the next 500 customers, 72 bought items recommended by the new model. → $\frac{72}{500} \approx 14\%$

Is the improvement brought by the new model is statistically significant at 95% confidence?

$\alpha = 0.05$

Soln. $H_0: p = 0.1$

$H_A: p > 0.1$

- Test statistic : $X \rightarrow$ # of people who bought the recommended items.

What distribution? $\rightarrow 0, 1, 2, \ldots, 500$ → Binomial ($n = 500$)

$X \sim \text{Binomial}(n = 500, p = 0.1)$

⇒ Right tailed Test ✓

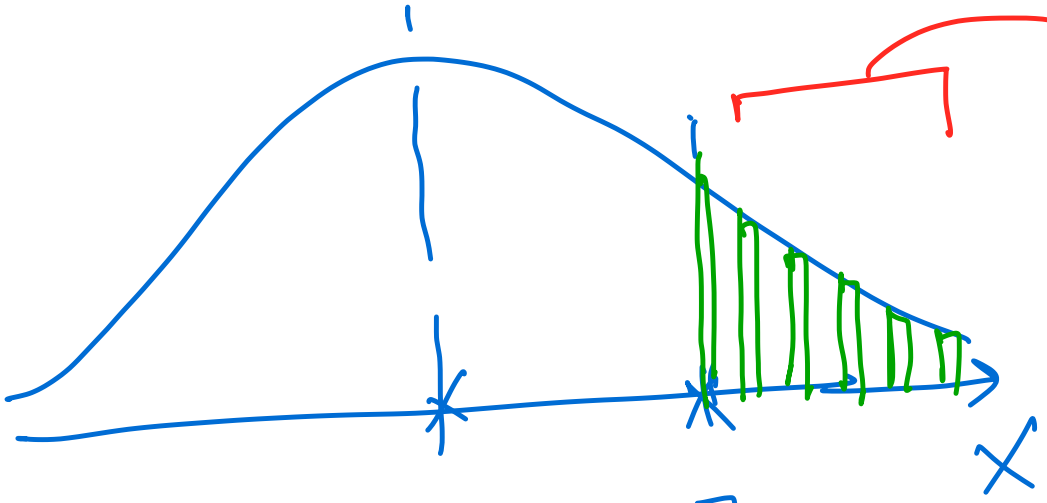
p-value:

$P[X \geq 72 \mid H_0 \text{ is True}] = p\text{-value}$

$= 1 - P[X \leq 71 \mid H_0 \text{ is True}]$

$= 1 - \text{Binom.cdf.}$

$(n = 500,\ R = 71,\ p = 0.1)$



$= 0.001$

$p\text{-value} = 0.001$

$\alpha = 0.05$

$p\text{-value} \ll \alpha$

Reject H_0