통계분석
Statistical Analysis

Testing Hypotheses : 가설 검증

Hypothesis

- **Hypothesis on what?**

  → Hypothesis on parameters (average, variance, etc.) of the population

Testing Hypothesis: Example

Testing unfair coin

$p =$ probability that the coin shows heads for single tossing

- $p = 0.5$: Fair coin
- $p \neq 0.5$: Unfair coin

- How can we verify whether the coin is unfair or not?

  Do tossing experiments, for example, tossing 20 times.

  Sample data

  X = the number of heads among 20 flips.

  Statistic

  Using this statistic, we will verify whether this coin is unfair.

Hypothesis: Example

Setting up Hypothesis to prove

The coin is fair, i.e., $p = 0.5$.

- If a statistic from 20 tossing experiments “statistically” shows the initial hypothesis is true, we accept this hypothesis.

- If we cannot find out that the hypothesis is true from experiments, we do not take the initial hypothesis.

  In this case, we take a hypothesis contradictory to the initial one.

  For example, we might take one hypothesis among $\begin{cases} p \neq 0.5 \\ p < 0.5 \\ p > 0.5 \end{cases}$

Experiments and Test Statistic

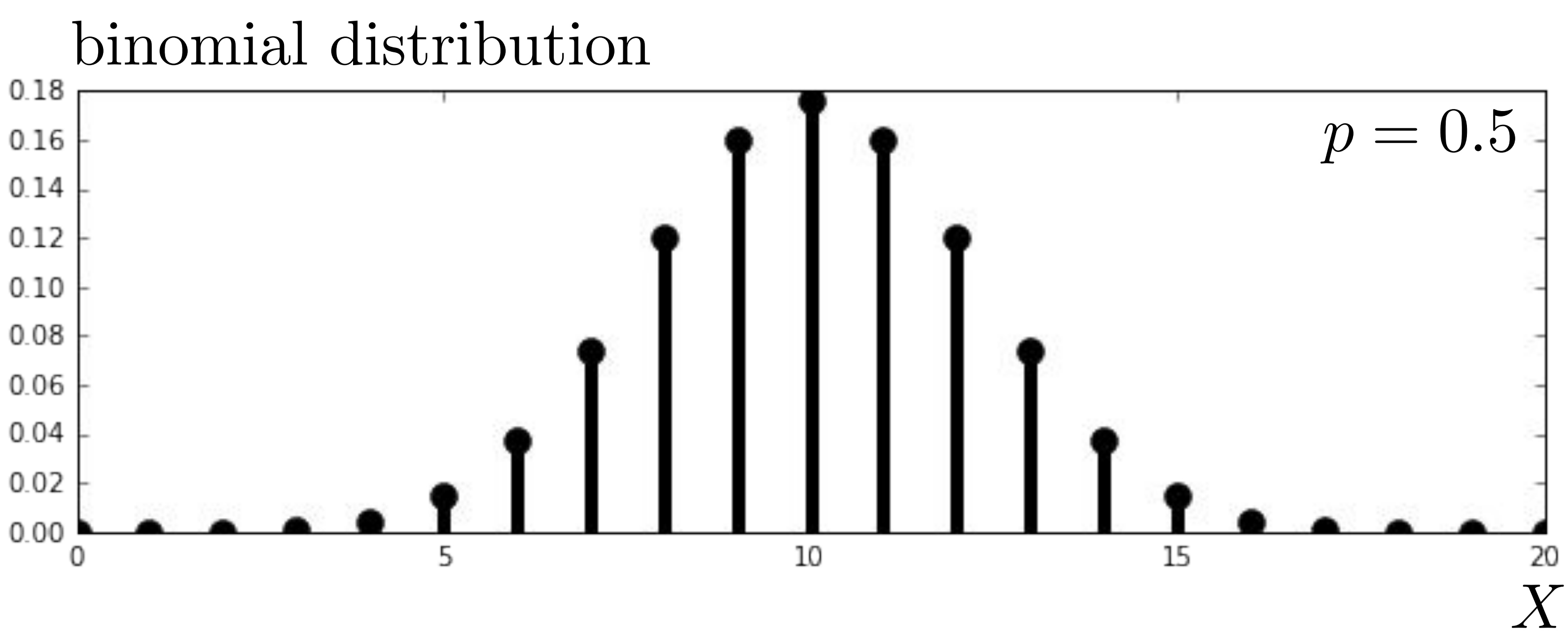
Experiments

Toss the coin 20 times, and count the number of heads among total trials.

Test statistic

$X =$ the number of heads among total trials.

- Test statistic X follows the binomial distribution.
- If our hypothesis is true, i.e., the coin is fair,



Setting up Hypotheses to test

- Null Hypothesis H_0 [귀무가설]

  Claim assumed to be true initially

- Alternative Hypothesis H_a [대립가설]

  Assertion that we take when the null hypothesis is rejected

- Alternative hypothesis is contradictory to Null one.
- It means that we CANNOT take both of them at the same time.

Hypotheses: Tossing a coin

- Null Hypothesis H_0 [귀무가설]

  Claim assumed to be true initially

  Let us say initially we believe that the coin is fair. $p = 0.5$

- Alternative Hypothesis H_a [대립가설]

  Assertion that we take when the null hypothesis is rejected

  Possible alternative hypotheses $\begin{cases} p \neq 0.5 \\ p < 0.5 \\ p > 0.5 \end{cases}$

Simple / Composite Hypothesis

- Null Hypothesis H_0 [귀무가설]

  Claim assumed to be true initially

  Simple Hypothesis

  Let us say initially we believe that the coin is fair. $p = 0.5$

- Alternative Hypothesis H_a [대립가설]

  Assertion that we take when the null hypothesis is rejected

  Possible alternative hypotheses $\left\{ \begin{array}{l} p \neq 0.5 \\ p < 0.5 \\ p > 0.5 \end{array} \right.$

  Composite Hypothesis

- Simple Hypothesis: Hypothesis whose distribution we can specify

For example, Simple hypothesis that a parameter of the distribution has some specific value.

- Composite Hypothesis: Hypothesis whose distribution we cannot specify

A parameter of the distribution lies in some range.

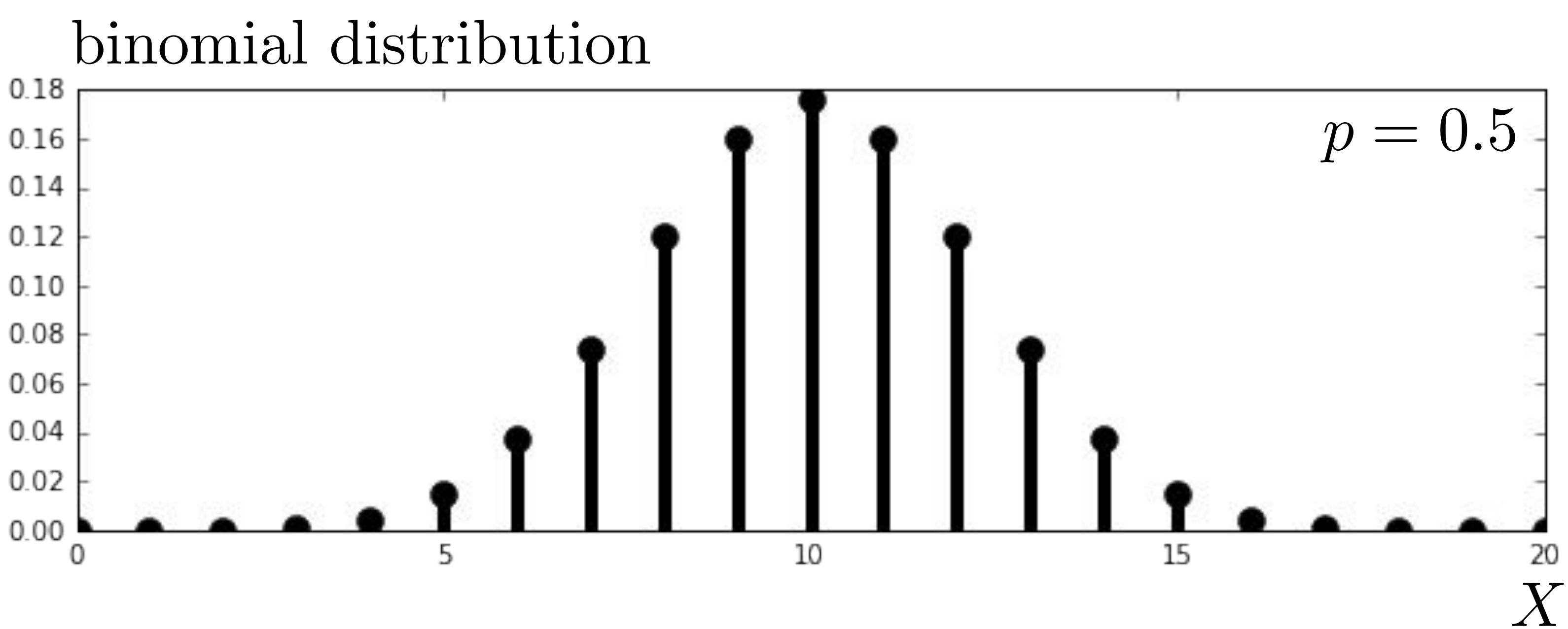
Test Statistic on Hypothesis Testing

- Test Statistic X

  A function of sample data which we use to make a decision

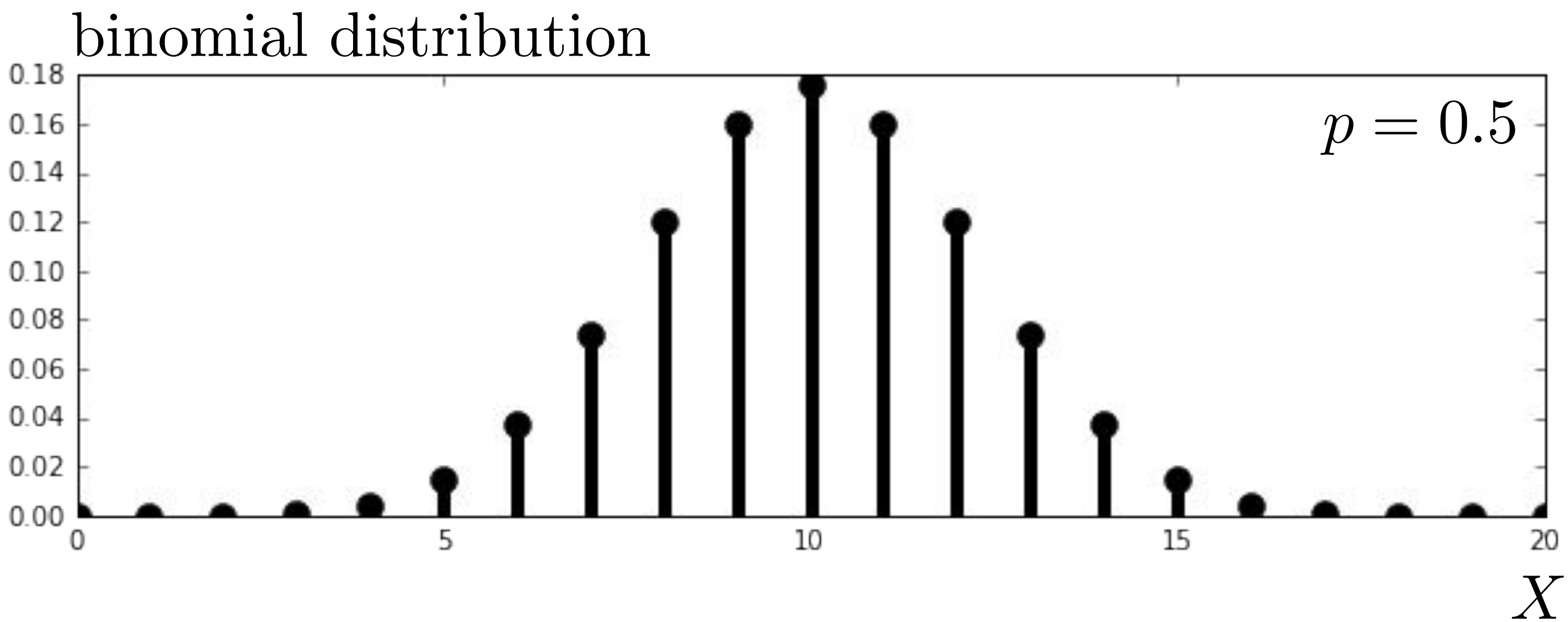
  $X =$ the number of heads among total trials.

- Test statistic X follows the binomial distribution.
- If the null hypothesis is true, i.e., the coin is fair,



Binomial Distribution: Null Hypothesis

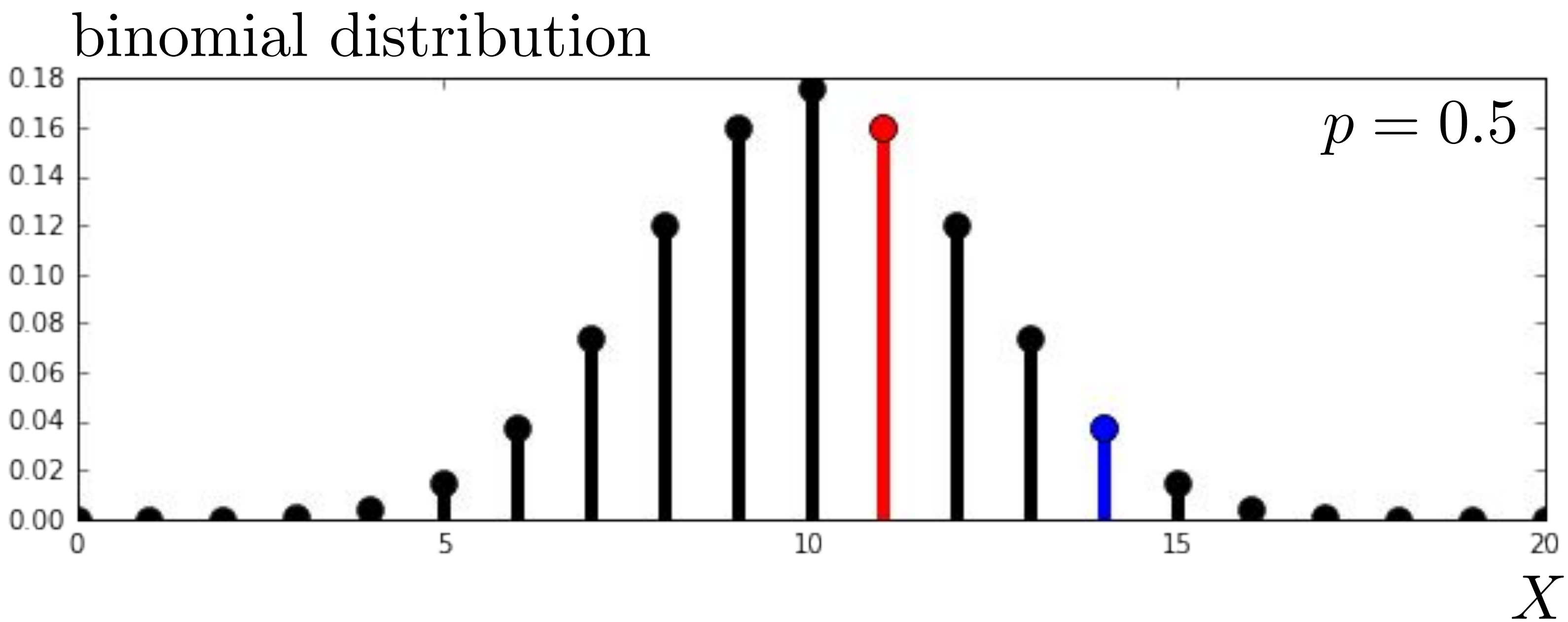
- Test Statistic $X =$ the number of heads among 20 trials.



When $p = 0.5$ $\begin{cases} \text{X is not always 10.} \\ X = 10 \text{ is the most probable.} \\ \text{It is also likely that X has a value close to 10.} \end{cases}$

Binomial Distribution: Null Hypothesis

- Test Statistic $X =$ the number of heads among 20 trials.



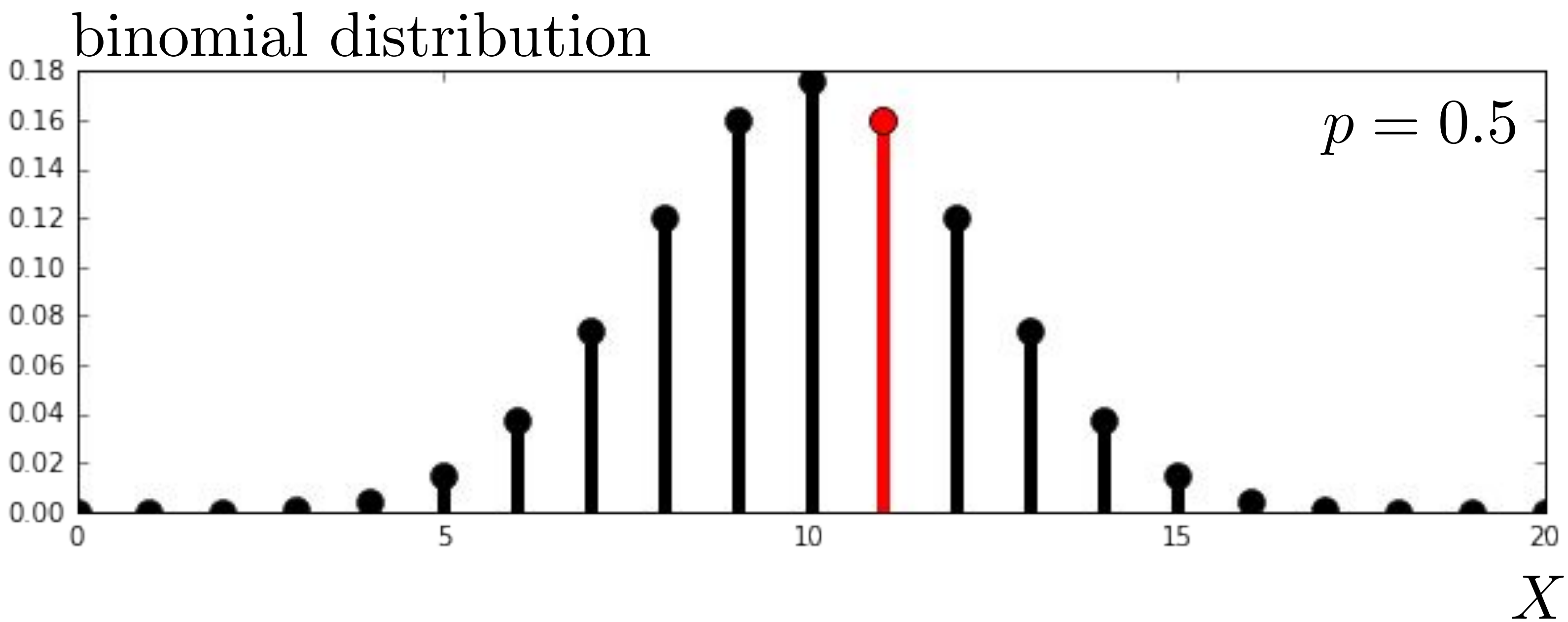
When $p = 0.5$ $\left\{ \begin{array}{l} \text{for example, } X = 11 \text{ is likely to happen.} \\ X = 14 \text{ is less probable than } X = 11, \\ \text{but } X = 14 \text{ still has a nonzero probability.} \end{array} \right.$

Test Procedure

- Null hypothesis $H_0 : p = 0.5$
- Alternative hypothesis $H_0 : p \neq 0.5$
- Test Statistic $X =$ the number of heads among 20 trials.

Null Hypothesis Accepted

- Test Statistic $X =$ the number of heads among 20 trials.



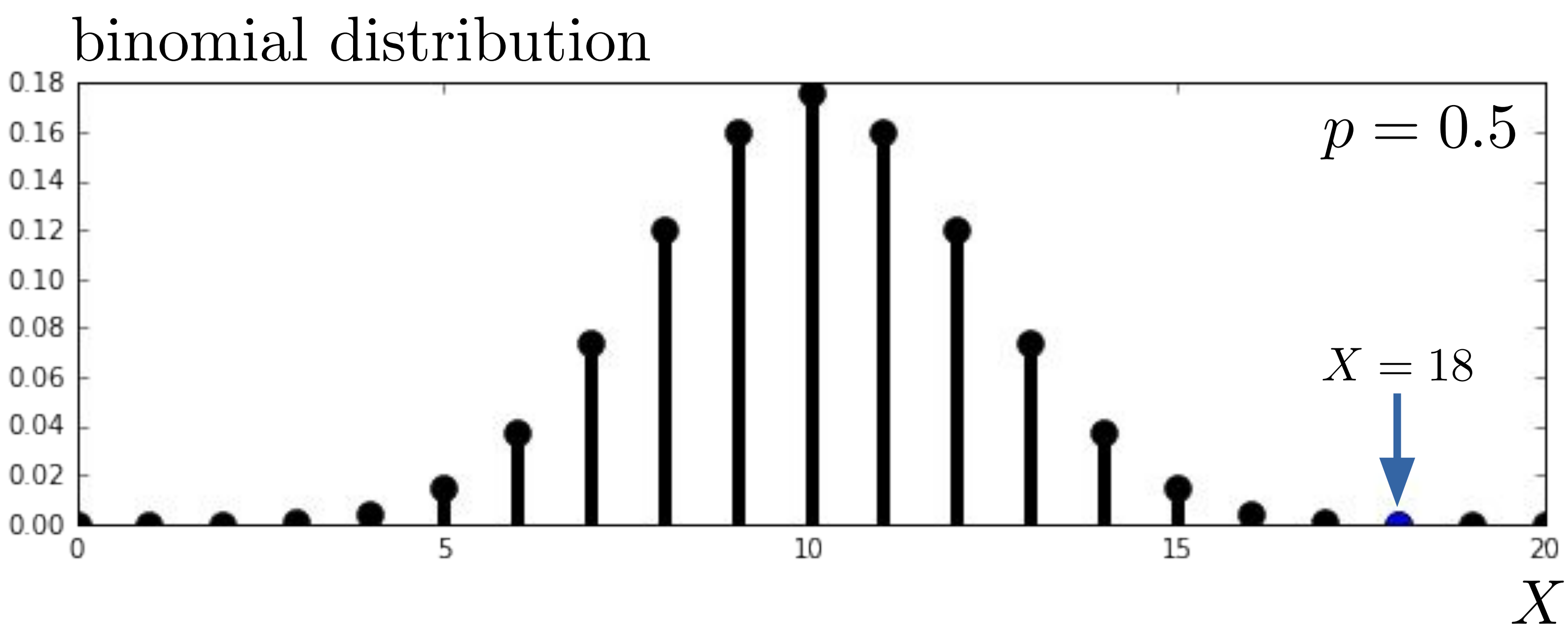
Let us say that $X = 11$ in our experiment.

Then, we might think $p = .5$ is a reasonable hypothesis.

"Null Hypothesis is accepted."

Null Hypothesis Rejected

- Test Statistic $X =$ the number of heads among 20 trials.

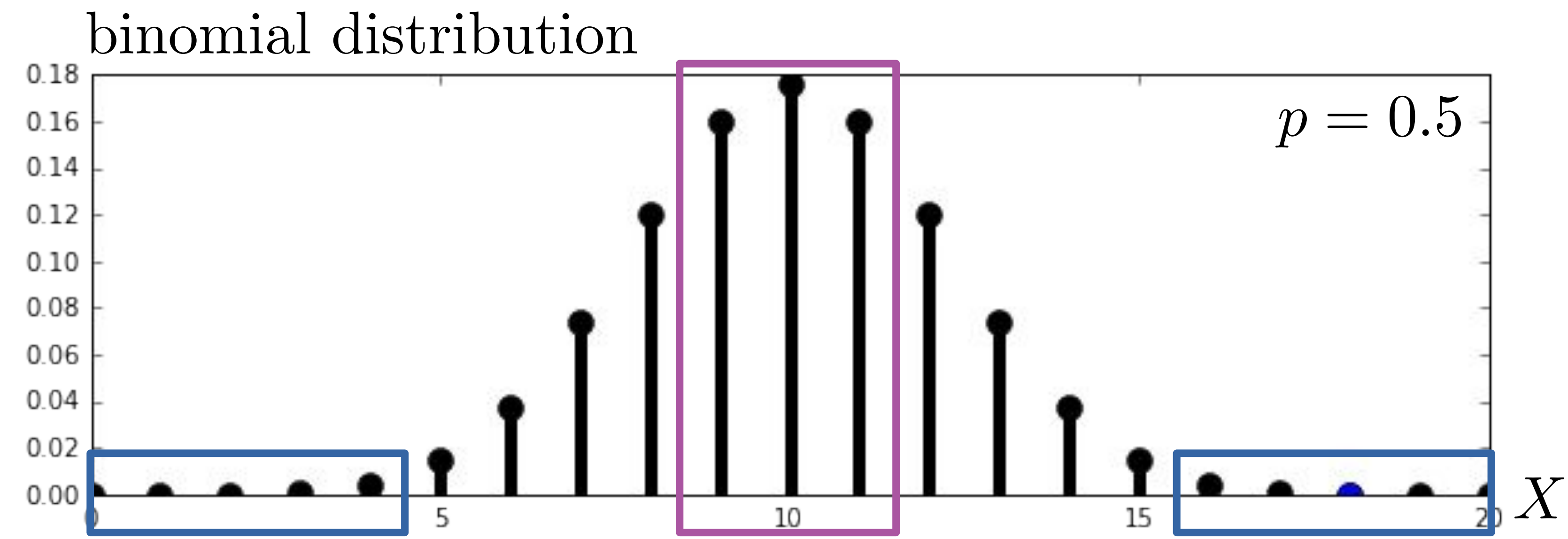


Let us say that $X = 18$ in our experiment.
It is pretty much unlikely that $p = 0.5$.

Null Hypothesis is rejected.
Instead, alternative hypothesis is favored.

When Accepted or Rejected?

- Test Statistic $X =$ the number of heads among 20 trials.



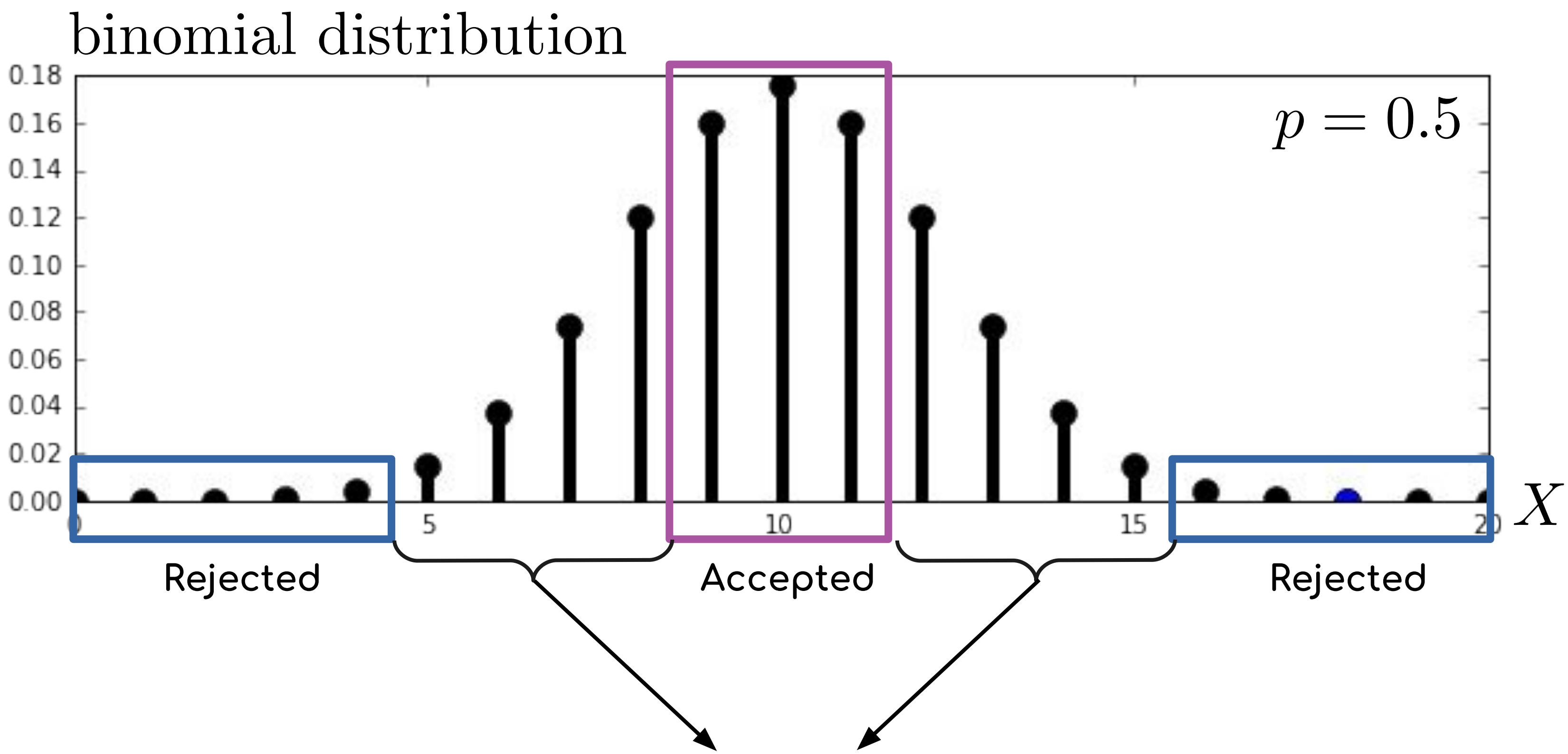
If X is included in this region, the null hypothesis rejected

If X is included in this region, the null hypothesis accepted

If X is included in this region, the null hypothesis rejected

When Accepted or Rejected?

- Test Statistic $X =$ the number of heads among 20 trials.

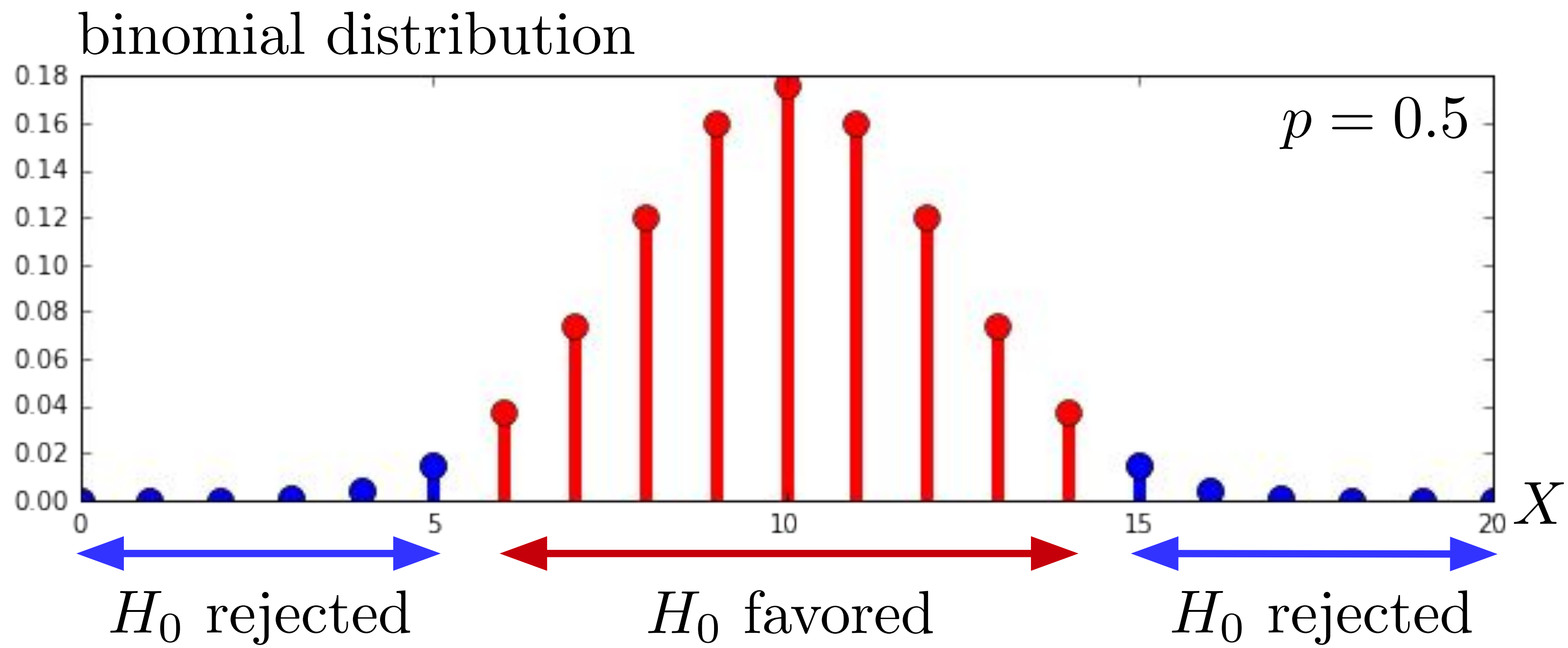


If X is in these ranges, the null hypothesis is accepted or rejected??

→ We have to decide the boundaries where the null hypothesis is accepted or rejected.

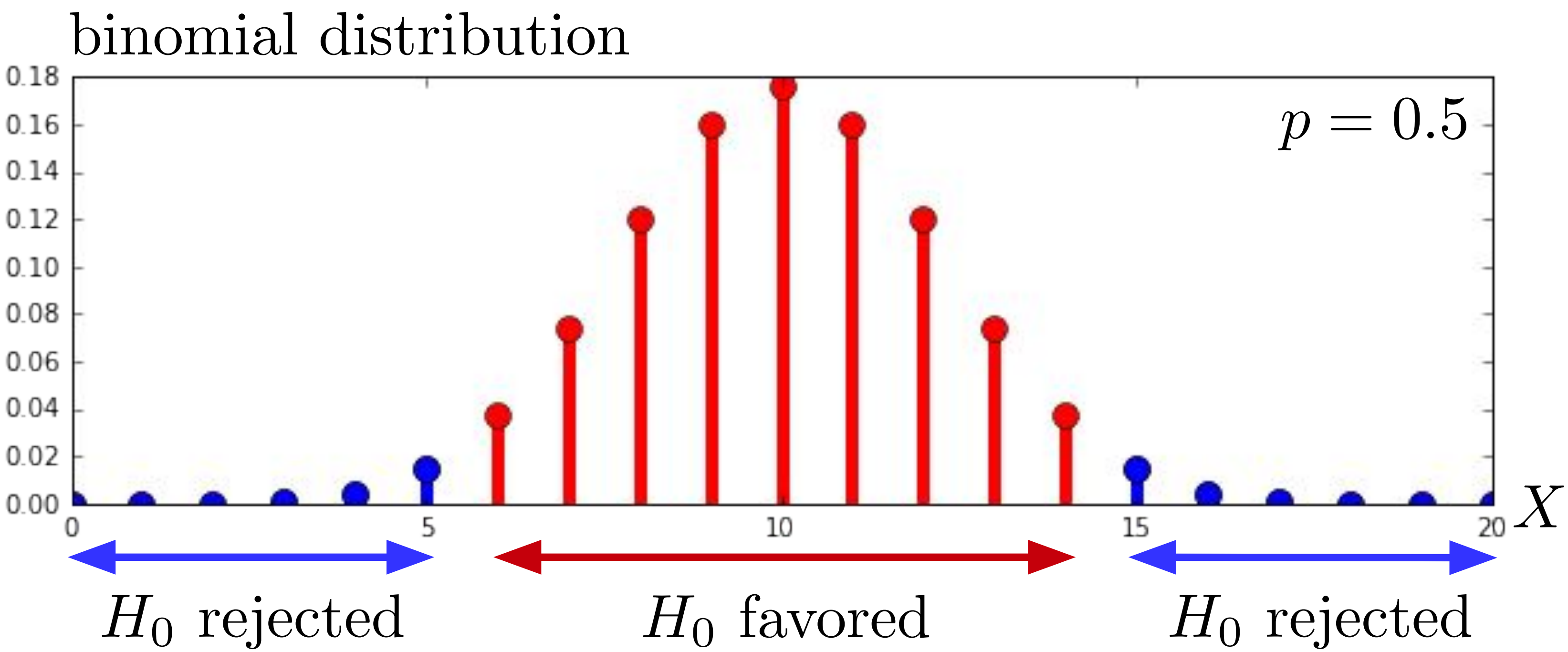
Rejection Region

- Test Statistic $X =$ the number of heads among 20 trials.



- Rejection region = A set of test statistics which implies the null hypothesis is rejected.
- Here, Rejection region = {0,1,2,3,4,5,15,16,17,18,19,20}

Statistical Errors I

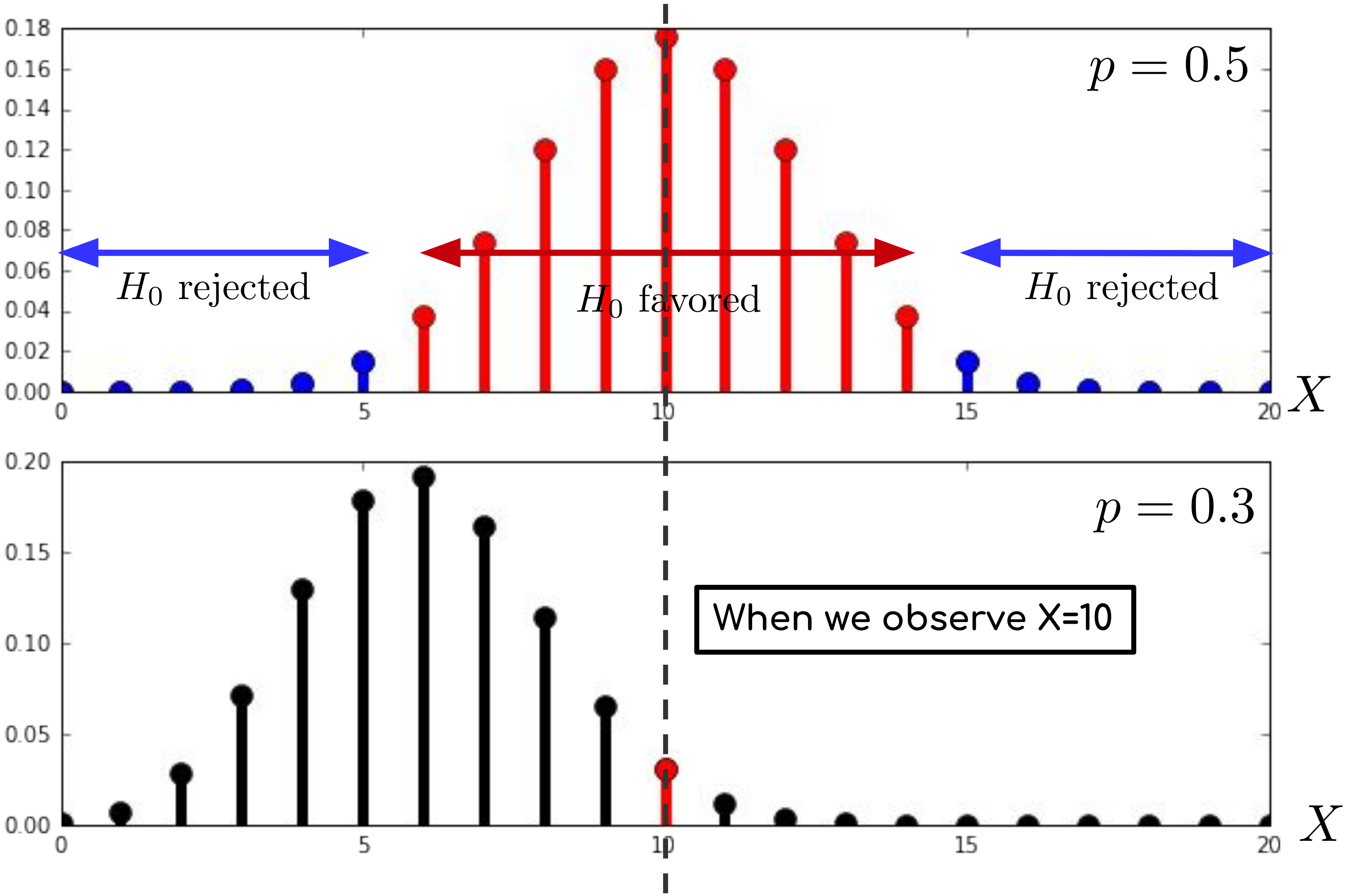


HOWEVER!

- Even if we have a test statistic included in our rejection region, it does NOT mean that the null hypothesis is absolutely wrong.
- There is still (small but) nonzero probability that a test statistic is in the rejection region, when the null hypothesis is true.

Statistical Errors II

It is also possible that a test statistic is not in the rejection region, when the null hypothesis is false.



When $p = 0.3$, there is 0.0308 probability for $X = 10$.

Type I and II Errors

Decision		Truth: $H_0 = \text{True}$	$H_a = \text{True}$
	Reject H_0	Type I Error	Right decision
	Not Reject H_0	Right decision	Type II Error

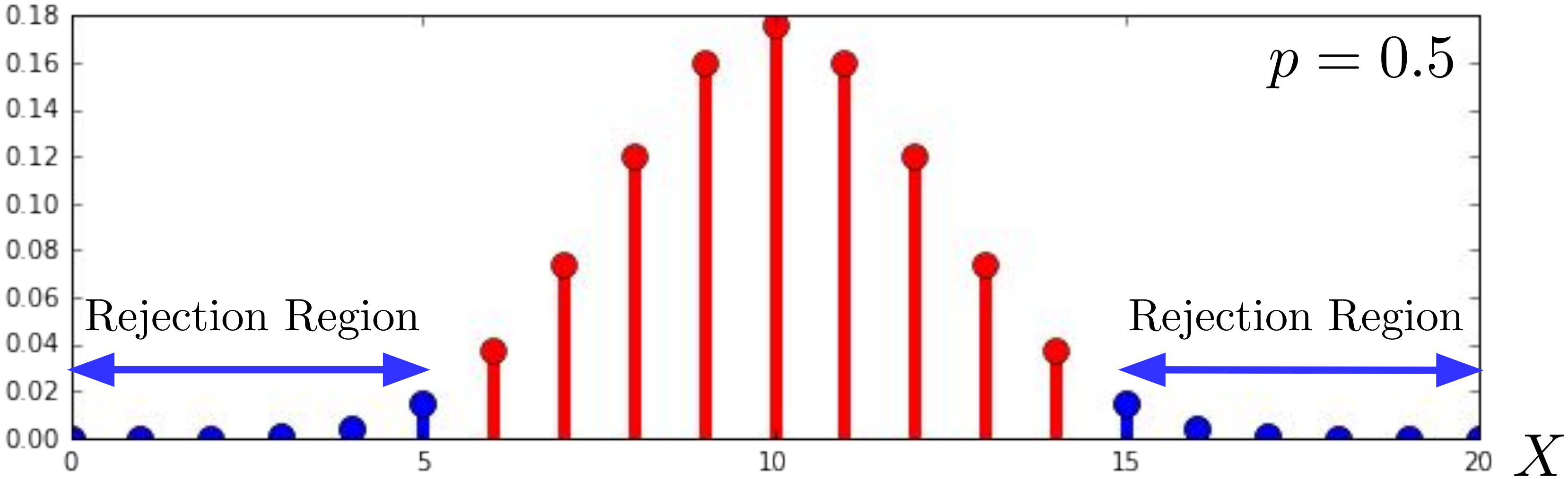
- Type I Error: the null hypothesis is rejected when it is true.
- Type II Error: the null hypothesis is not rejected when it is false.

$\rightarrow$ We need to quantify these errors.
We need to express these errors in terms of probability language.

Type I Error

- Test Statistic $X =$ the number of heads among 20 trials.

binomial distribution



$$\text{Type I Error} = P(\text{Reject } H_0 | H_0 \text{ is true})$$

Conditional probability

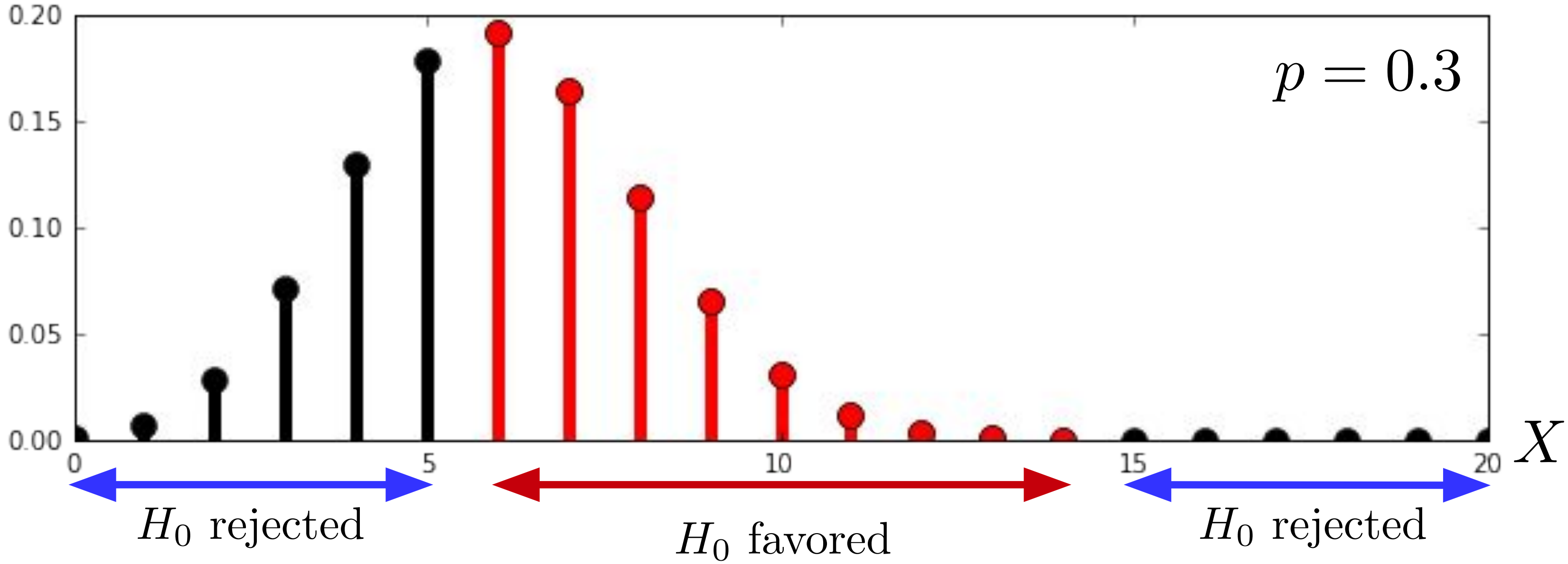
$$P(X \in \{\text{Rejection region}\} | p = 0.5) \approx 0.0414$$

Type II Error

$$\text{Type II Error} = P(\text{Not reject } H_0 | H_0 \text{ is false})$$

- Unlike Type I Error, Type II error is not a single value, because the underlying distribution is not specified.
- After specifying the underlying distribution supporting the alternative hypothesis, one can calculate Type II error.

Type II Error

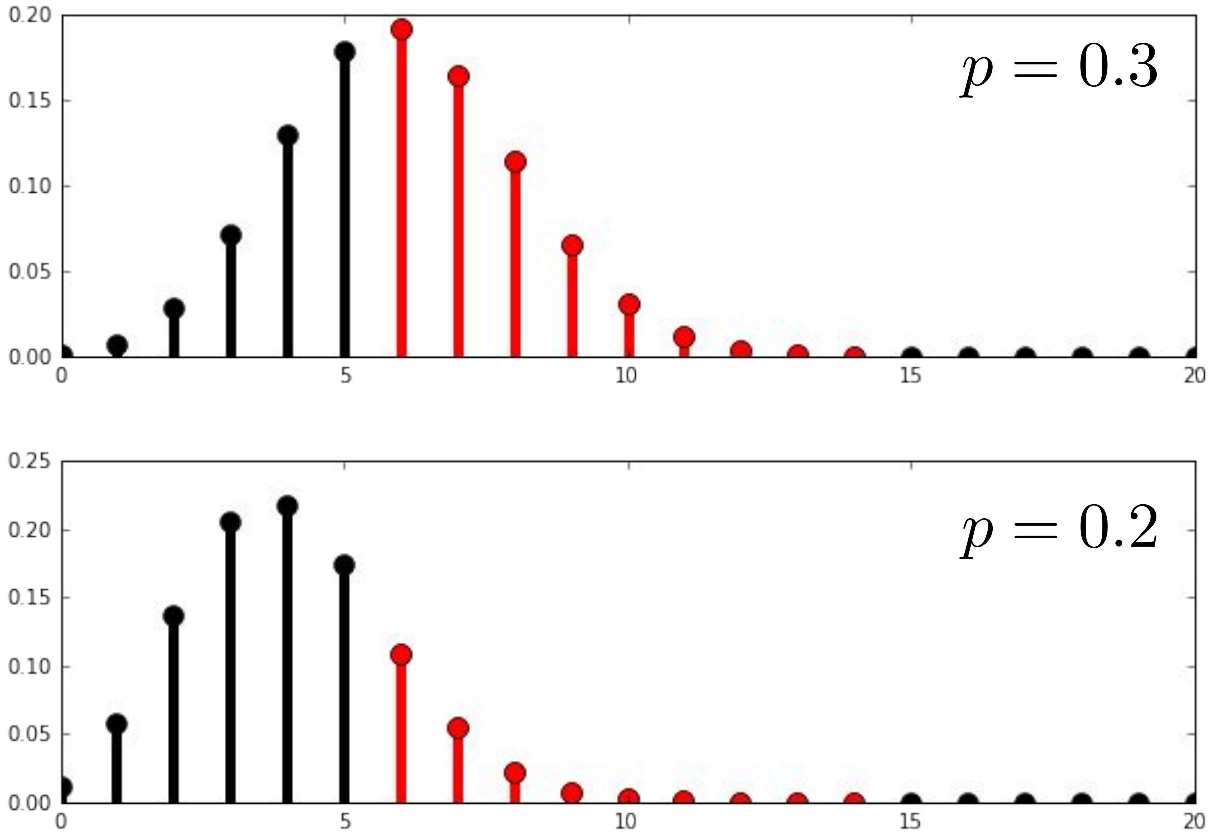


$$\text{Type II Error} = P(\text{Not reject } H_0 | H_0 \text{ is false})$$

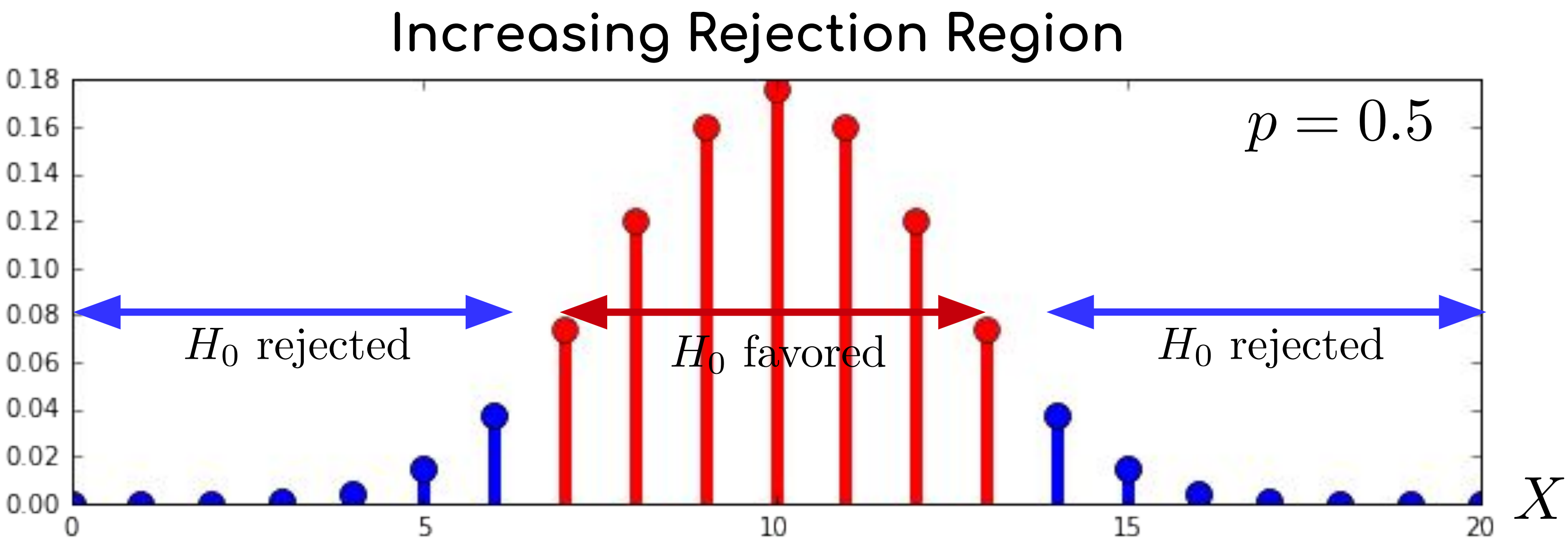
$$\beta(p = .3) = P(X \notin \{\text{Rejection region}\} | p = 0.3) \approx 0.5836$$

Type II Error

$$\beta(p = .1) = \beta(p = .9) = 0.0113$$
$$\beta(p = .2) = \beta(p = .8) = 0.1958$$
$$\beta(p = .3) = \beta(p = .7) = 0.5836$$
$$\beta(p = .4) = \beta(p = .6) = 0.8728$$



Changing Rejection Region

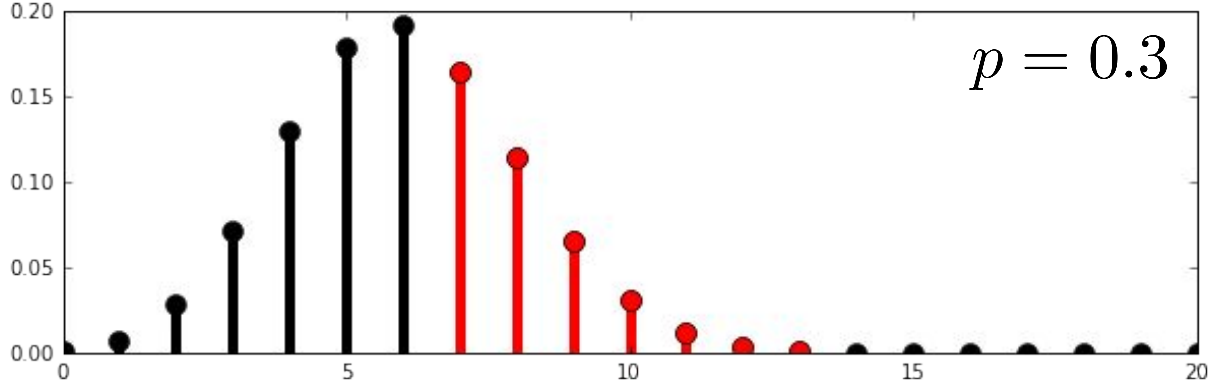


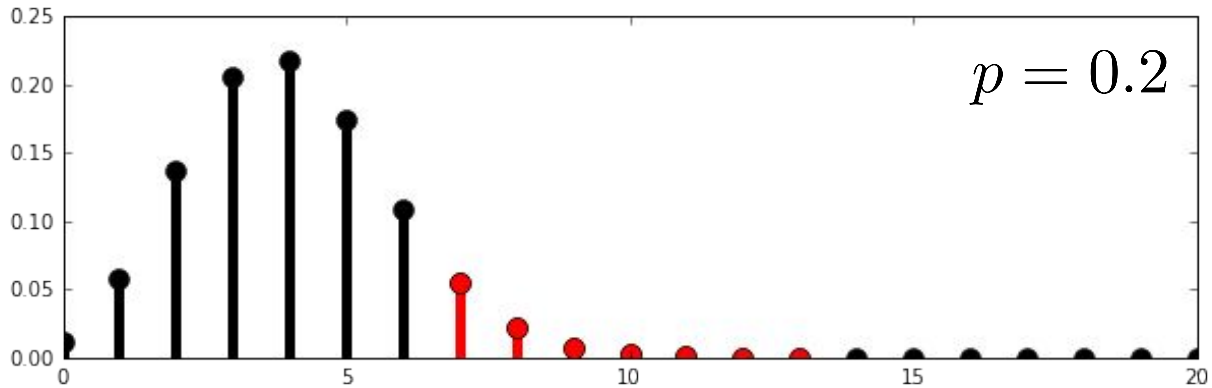
$$\text{Type I Error} = P(\text{Reject } H_0 | H_0 \text{ is true})$$

$$P(X \in \{\text{Rejection region}\} | p = 0.5) \approx 0.1153$$

Changing Rejection Region

$$\beta(p = .1) = \beta(p = .9) = \ 0.00239$$
$$\beta(p = .2) = \beta(p = .8) = \ 0.0867$$
$$\beta(p = .3) = \beta(p = .7) = \ 0.3917$$
$$\beta(p = .4) = \beta(p = .6) = \ 0.7435$$





Changing Rejection Region

- **Rejection Region = {0,1,2,3,4,5,15,16,17,18,19,20}**

Type I Error $$P(X \in \{\text{Rejection region}\} | p = 0.5) \approx 0.0414$$

Type II Error
$$\beta(p = .1) = \beta(p = .9) = 0.0113$$
$$\beta(p = .2) = \beta(p = .8) = 0.1958$$
$$\beta(p = .3) = \beta(p = .7) = 0.5836$$
$$\beta(p = .4) = \beta(p = .6) = 0.8728$$

- **Rejection Region = {0,1,2,3,4,5,6,14,15,16,17,18,19,20}**

Type I Error $$P(X \in \{\text{Rejection region}\} | p = 0.5) \approx 0.1153$$

Type II Error
$$\beta(p = .1) = \beta(p = .9) = 0.00239$$
$$\beta(p = .2) = \beta(p = .8) = 0.0867$$
$$\beta(p = .3) = \beta(p = .7) = 0.3917$$
$$\beta(p = .4) = \beta(p = .6) = 0.7435$$

Trade-off between Type I and II Errors

- If we increase the rejection region, then Type I Error increases, but Type II Error decreases.
- If we reduce the rejection region, then Type I Error decreases, but Type II Error increases.