The Hypothesis test procedure:

1. Define H_0 and H_1
2. Assuming H_0 to be true
3. Compute a test statistic to assess the strength of the evidence against H_0
4. Compute the critical region or p-value
5. State a conclusion about the strength of the evidence against H_0.

Example: Highway engineers have found that the proper alignment of the vehicle's headlights is a major factor affecting the performance of reflective highway signs. It is thought that more than 50% of the vehicle on the road have misaimed headlights. If this contention can be supported statistically, then a new tougher inspection program will be put into operation.

Let p denote the proportion of vehicle in operation that have misaimed headlights. Since we wish to support the statement that $p > 0.5$, we have the two hypothesis:

$$H_0: p \leq 0.5$$
$$H_1: p > 0.5$$

A random sample of 20 cars is selected and the head lights are tested. Let X be the number of vehicles in the sample with misaimed headlights. X is a binomial random variable. Then, assuming H_0 is true ($p = 0.5$), the average number of vehicle with misaimed headlights should be

$$np = 20 \times 0.5 = 10$$

If the value of X is high, say $X > 14$, we tend to reject H_0, i.e., the sampled data support that $p > 0.5$

Otherwise, H_0 will not be rejected. i.e., we fail to reject H_0

X is the test statistic and a critical value (14 in this example) is calculated.

Two types of errors

Possible results for any test of a Hypothesis:

1. We rejected H_0 when it was true, this is Type I error
2. We rejected H_0 when H_1 was true, correct decision
3. We failed to reject H_0 when H_1 was true, this is Type II error
4. We failed to reject H_0 when H_0 was true, correct decision

	Null Hypothesis (H0)	
Decisions	*True*	*False*
Fail to reject H0	Correct decision	***Type II error***
Reject H0	***Type I error***	Correct decision

In the Highway engineering example, we were testing $H_0: p \leq 0.5$, $H_1: p > 0.5$
If we reject H_0 when H_0 is true, a Type I error is made. This error could lead to the implementation of an unnecessary inspection program. (false positive rate, false alarm rate)
If we failed to reject H_0 when H_1 is true, a Type II error is made. In the example, we concluded that the new inspection program is not necessary when in fact it is needed. (false negative rate)

In the lung cancer treatment medicine example, we were testing $H_0: p \leq 0.5$, $H_1: p > 0.5$
If we reject H_0 when H_0 is true, a Type I error is made. This error could mislead on the cure rate of the medicine and cause unnecessary delay on proper treatment of patients. (false positive rate, false alarm rate)
If we failed to reject H_0 when H_1 is true, a Type II error is made. This error could kill a good lung cancer treatment medicine and cause a tremendous financial loss of the developer. (false negative rate)

Note: any time H_0 is rejected, a Type I error may occur; any time H_0 is not rejected, a Type II error may occur;
There is no way to avoid both.

Hypothesis Test on parameter $\boldsymbol{\theta}$

The Critical Region method:

The purpose of hypothesis test is to decide whether the evidence tend to reject the null hypothesis. This involve a procedure to find the critical (or rejection) region for the test. ***The critical region constitute the values of the test statistic that lead to rejection of the null hypothesis.***

Significance of the test:

The probability that the observed value of the test statistic will fall into the rejection region by chance even though $\theta = \theta_0$ (H_0 is true) is called α, the *level of significance* of the test. α is the probability of committing a Type I error.

Example: Highway engineers have found that the proper alignment of the vehicle's headlights is a major factor affecting the performance of reflective highway signs. It is thought that more than 50% of the vehicle on the road have misaimed headlights. If this contention can be supported statistically, then a new tougher inspection program will be put into operation.
we have the two hypothesis:

$$H_0: p \leq 0.5;\ H_1: p > 0.5$$

A random sample of 20 cars is selected and the head lights are tested.

Let X be the number of cars in the sample with misaimed headlights. X is the test statistic. (***choose the test statistic with know distribution***)

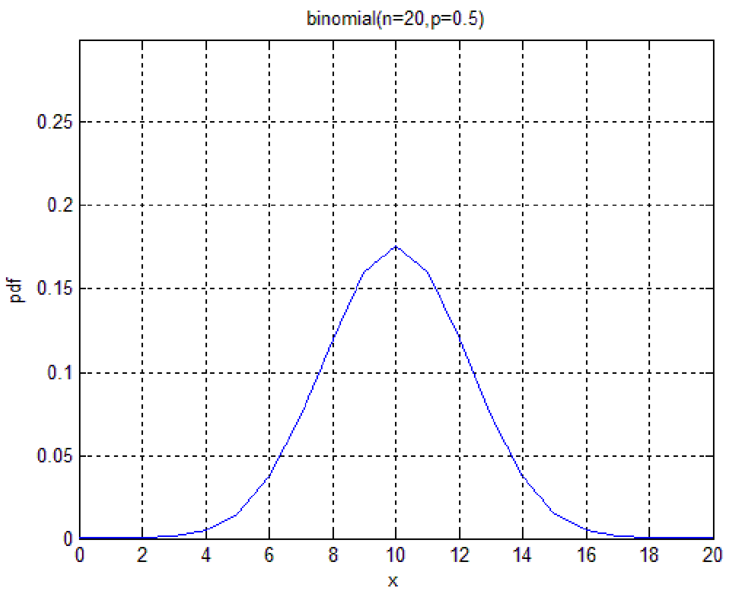
If H_0 is true ($p \leq 0.5$), then X is binomial with $n = 20, p = 0.5$ $(?)$.

$$E[X] = np = 10$$

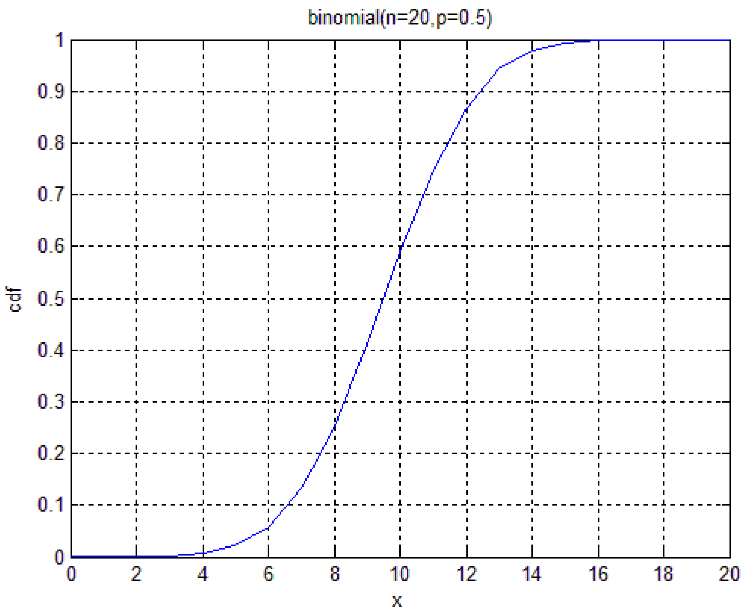
If the observed value of X is higher than a critical value x_0, we tend to reject H_0.

How to find the value of this critical value of X, x_0?

What is the consequence if x_0 is chosen too small?



Example: (continued)



This critical value should be chosen in such a way that the probability that we make Type I error will not exceed α. (The probability that H_0 is rejected when actually H_0 is true).

According to this definition, we have,

$$P[X \geq x_0 | p = 0.5] \approx \alpha$$

i.e.,

$$1 - P[X < x_0 | p = 0.5] \approx \alpha \Rightarrow P[X < x_0 | p = 0.5] \approx 1 - \alpha$$

From the accumulative probability table of binomial, we have, ($\alpha = 0.05$)

$$P[X \leq 13 | p = 0.5] = 0.942 \Rightarrow P[X < 14 | p = 0.5] = 0.942$$

i.e.,

$$P[X \geq 14 | p = 0.5] = 1 - 0.942 = 0.058$$

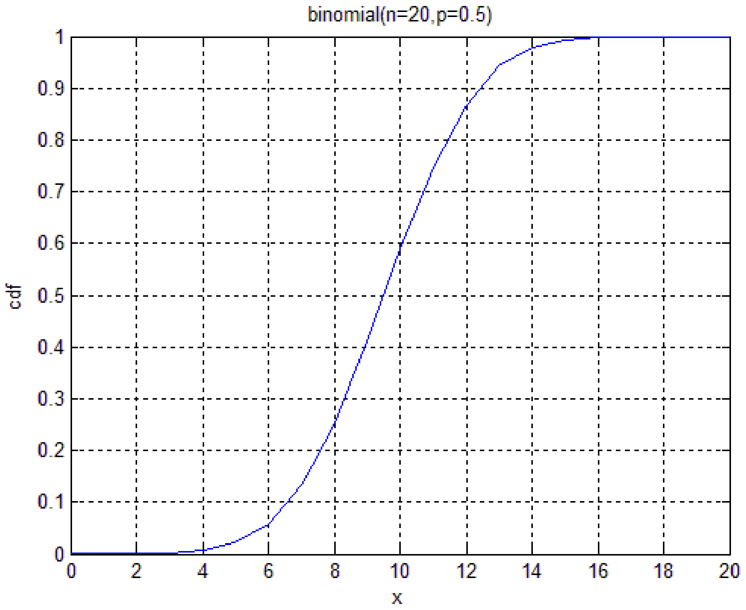
Then, we reject H_0 if the observed value of X is 14 or greater. The critical (or rejection) region is $c = \{14,15,16,17,18,19,20\}$. *This is a right-tailed region on the distribution*.

Question: Why I can use $p = 0.5$ to represent $H_0: p \leq 0.5$?

Any value of X that lead us to reject 0.5 as a value of p also leads us to reject any value less than 0.5

In previous example, the critical (or rejection) region is found as $c = \{14,15,16,17,18,19,20\}$. When the significance level of the test is set as $\alpha = 0.05$.

Q: How will the critical region change if the significance level of the test is changed to $\alpha = 0.01$?



Assuming H_0 is true, i.e., $p = 0.5$, we want to find the critical value x_0 so that the probability we make type I error will not exceed $\alpha = 0.01$. This is:

$$P[X \geq x_0 | p = 0.5] \approx 0.01$$

$$\Rightarrow 1 - P[X < x_0 | p = 0.5] \approx 0.01 \Rightarrow P[X < x_0 | p = 0.5] \approx 0.99$$

From the accumulative probability table of binomial, we have, $(\alpha = 0.01)$

$$P[X \leq 15 | p = 0.5] = 0.994 \Rightarrow P[X < 16 | p = 0.5] = 0.994$$

i.e.,

$$P[X \geq 16 | p = 0.5] = 1 - 0.994 = 0.006$$

Hence, the rejection region becomes $c = [16,17,18,19,20]$ when $\alpha = 0.01$

CUMULATIVE BINOMIAL PROBABILITIES

Tabulated values are P(<= k)
(Computations are rounded at the third decimal place.)

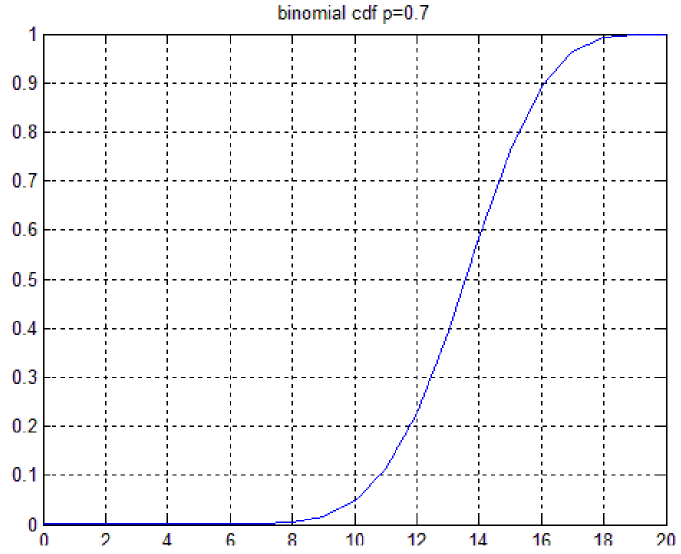
N = 20

k \ p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	0.122	0.012	0.001	0.000	0.000	0.000	0.000	0.000	0.000
1	0.392	0.069	0.008	0.001	0.000	0.000	0.000	0.000	0.000
2	0.677	0.206	0.035	0.004	0.000	0.000	0.000	0.000	0.000
3	0.867	0.411	0.107	0.016	0.001	0.000	0.000	0.000	0.000
4	0.957	0.630	0.238	0.051	0.006	0.000	0.000	0.000	0.000
5	0.989	0.804	0.416	0.126	0.021	0.002	0.000	0.000	0.000
6	0.998	0.913	0.608	0.250	0.058	0.006	0.000	0.000	0.000
7	1.000	0.968	0.772	0.416	0.132	0.021	0.001	0.000	0.000
8	1.000	0.990	0.887	0.596	0.252	0.057	0.005	0.000	0.000
9	1.000	0.997	0.952	0.755	0.412	0.128	0.017	0.001	0.000
10	1.000	0.999	0.983	0.872	0.588	0.245	0.048	0.003	0.000
11	1.000	1.000	0.995	0.943	0.748	0.404	0.113	0.010	0.000
12	1.000	1.000	0.999	0.979	0.868	0.584	0.228	0.032	0.000
13	1.000	1.000	1.000	0.994	0.942	0.750	0.392	0.087	0.002
14	1.000	1.000	1.000	0.998	0.979	0.874	0.584	0.196	0.011
15	1.000	1.000	1.000	1.000	0.994	0.949	0.762	0.370	0.043
16	1.000	1.000	1.000	1.000	0.999	0.984	0.893	0.589	0.133
17	1.000	1.000	1.000	1.000	1.000	0.996	0.965	0.794	0.323
18	1.000	1.000	1.000	1.000	1.000	0.999	0.992	0.931	0.608
19	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.988	0.878
20	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

k: Number of success
p: Probability of success of one trial

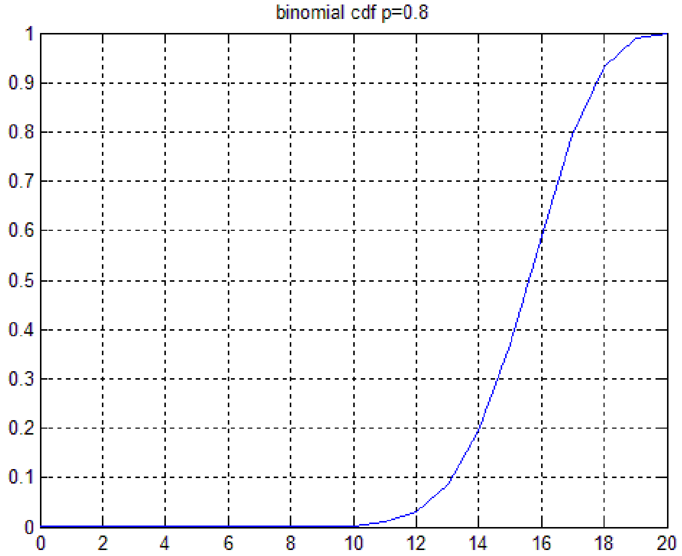
Type II error and β:

In hypothesis test, a type II error is an error that is made when the null hypothesis H_0 is not rejected when, in fact, the research hypothesis H_1 is true. The probability of committing a type II error is denoted by β.



Example: In the previous example, we find the critical region $c = \{14,15,16,17,18,19,20\}$. Suppose that, the true proportion of cars with misaimed headlights is 0.7. Then, β, the probability that H_0 will not be rejected given that $p = 0.7$, is calculated as:

$$\beta = P[type\ II\ error] = P[fail\ to\ reject\ H_0|H_1\ is\ true]$$
$$= P[X\ is\ not\ in\ the\ critical\ region|p = 0.7]$$
$$= P[X \leq 13|p = 0.7] = 0.3920$$



If the true value of p is changed from 0.7 to 0.8, then,

$$\beta = P[X \leq 13|p = 0.8] = 0.0867$$

Note, as the difference between the null value of 0.5 and the alternative value of p increases, β decreases.

Example: In the previous example, we find the critical region $c = \{14,15,16,17,18,19,20\}$. Suppose that, the true proportion of cars with misaimed headlights is 0.7. Then, β, the probability that H_0 will not be rejected given that $p = 0.7$, is calculated as:

$$\beta = P[X \leq 13|p = 0.7] = 0.3920$$

If the true value of p is changed from 0.7 to 0.8, then,

$$\beta = P[X \leq 13|p = 0.8] = 0.0867$$

Let's change the value of α from 0.05 to 0.01. The rejection region becomes $X \geq 16$. Then the β value becomes:

$$\beta = P[X \leq 15|p = 0.7] = 0.7625$$
$$\beta = P[X \leq 15|p = 0.8] = 0.3704$$

Note, when α decreases, β value increases.

Let's change the value of α from 0.05 to 0.1. The rejection region becomes $X \geq 13$. Then the β value becomes:

$$\beta = P[X \leq 12|p = 0.7] = 0.2277$$
$$\beta = P[X \leq 12|p = 0.8] = 0.0321$$

Note, when α increases, β value decreases.