Does smoking habit have an effect on the likelihood of having premature infants?

H_0: There is no effect on premature infants and smoking status.

H_a: There is an effect . . .

$$H_0: P_{smoking} - P_{nonsmoking} = 0$$

$$H_a: P_{smoking} - P_{non} \neq 0$$

smoking

Let $X = \{ \underbrace{X_1, \ldots, X_n}_{\text{random variables.}} \}$.

Bernoulli variables.
(0, 1)
not premie, premie

non smoking

$Y = \{ Y_1, \ldots, Y_n \}$

random variables

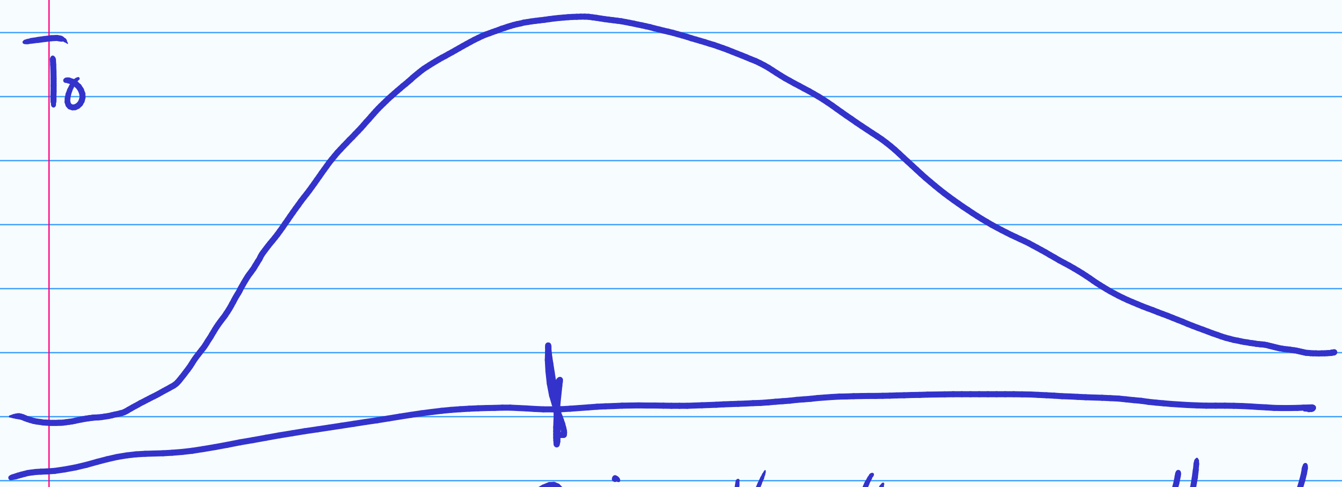
Bernoulli variables

(0 not premie, 1 premie)

$T_0(X, Y)$ = the test statistic based on our data, which incorporates our null hypothesis.

$T_0(X, Y)$ is a function of random variables. Therefore $T_0(X, Y)$ is a random variable.

$T_0 \sim F(\theta_0)$ $\quad$ θ_0 incorporates information about the null hypothesis.



θ_0: Hypothesized null value

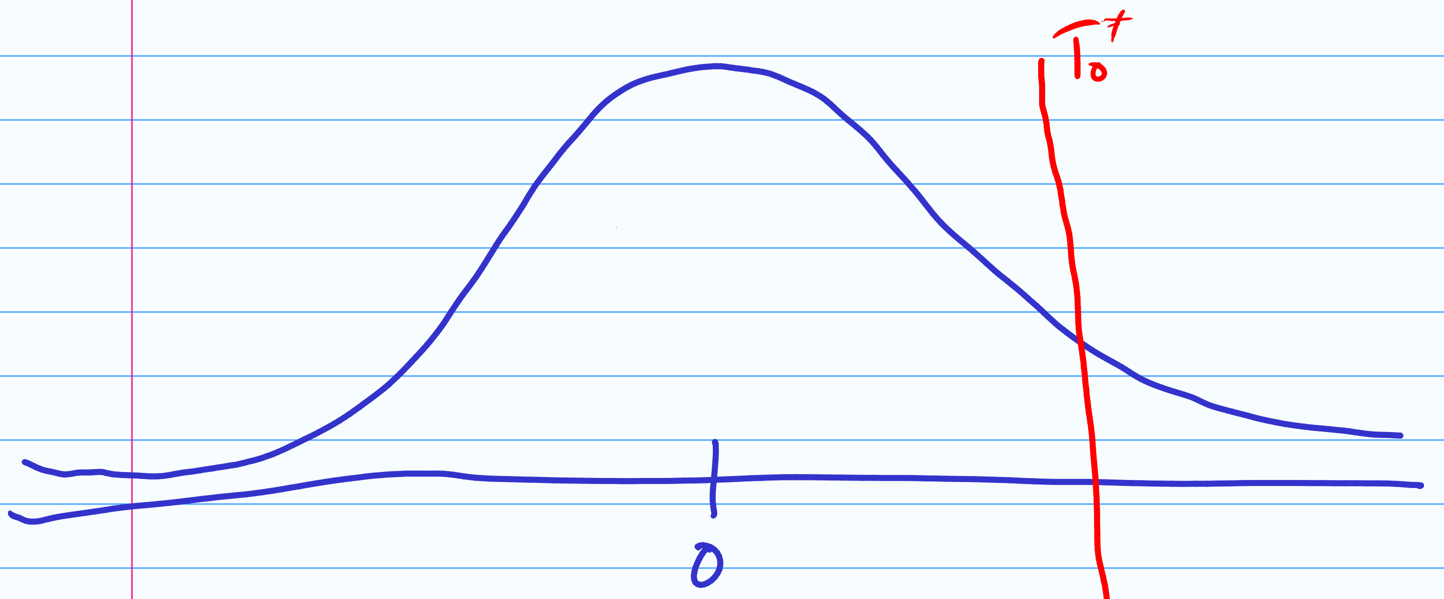
$\theta_0 = 0$ (0)

If we collect data

x^+ and y^+

We obtain a numerical value

$T_0(x^+, y^+) = T_0^+$ T_0^+



is T_0^+ close to 0?

or T_0^+ far away from 0?

Is there a cutoff point (T_{cv})

such that any value past T_{cv} is considered different from 0?

If so, how do we find it?

What is the value of T_{cv}?

$$P(T_0 > |T_{cv}|) = \alpha$$

predefined value between 0 and 1

We find the value of T_{cv} such that the probability is satisfied.

$$P(T_0 > T_{cv}) + P(T_0 < -T_{cv}) = \alpha$$

If the distribution is symmetrical, then

$$2 \cdot P(T_0 > T_{cv}) = \alpha$$

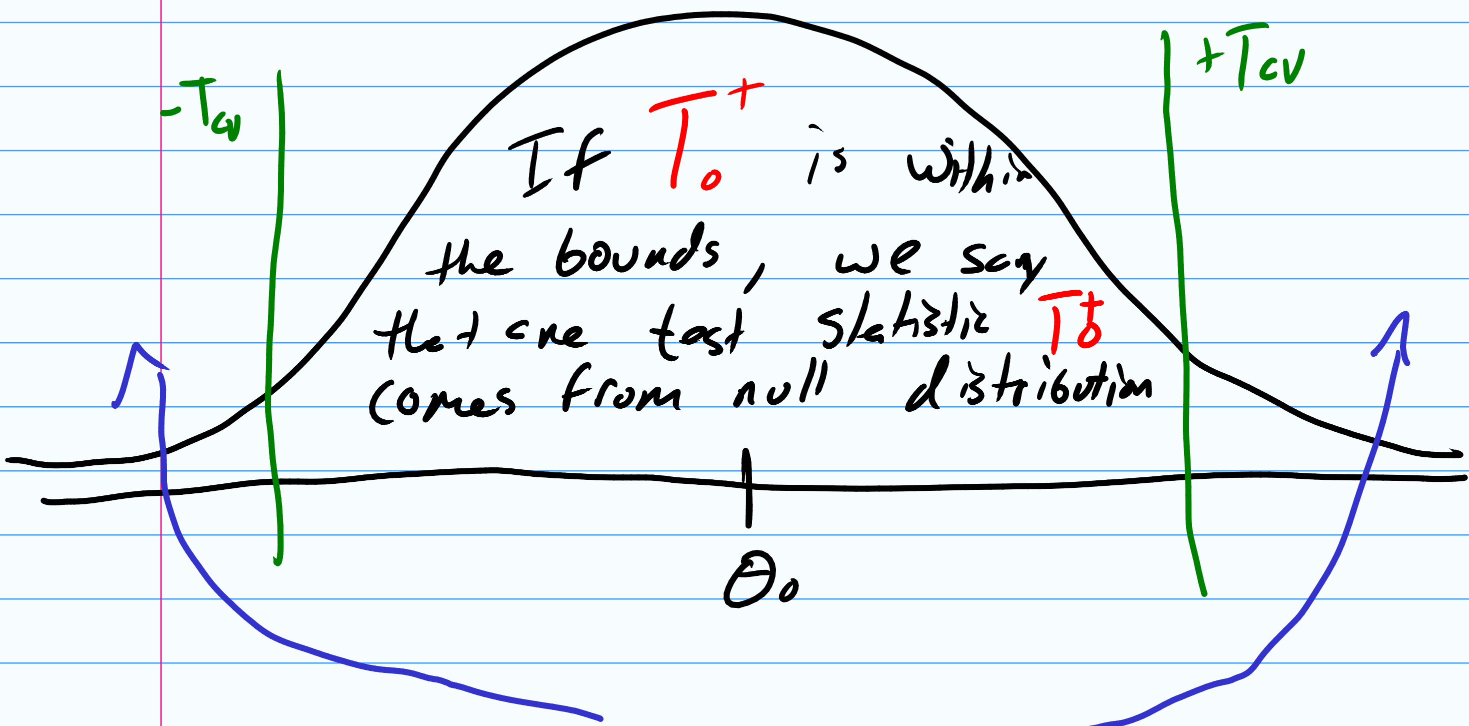
choices in α

$\alpha = 0.05$

For Biomedical studies

$\alpha = 0.01$

Once we obtain T_{cv}, we compare T_0^+ to the critical value and make a decision on whether the data comes from the null distribution or not.

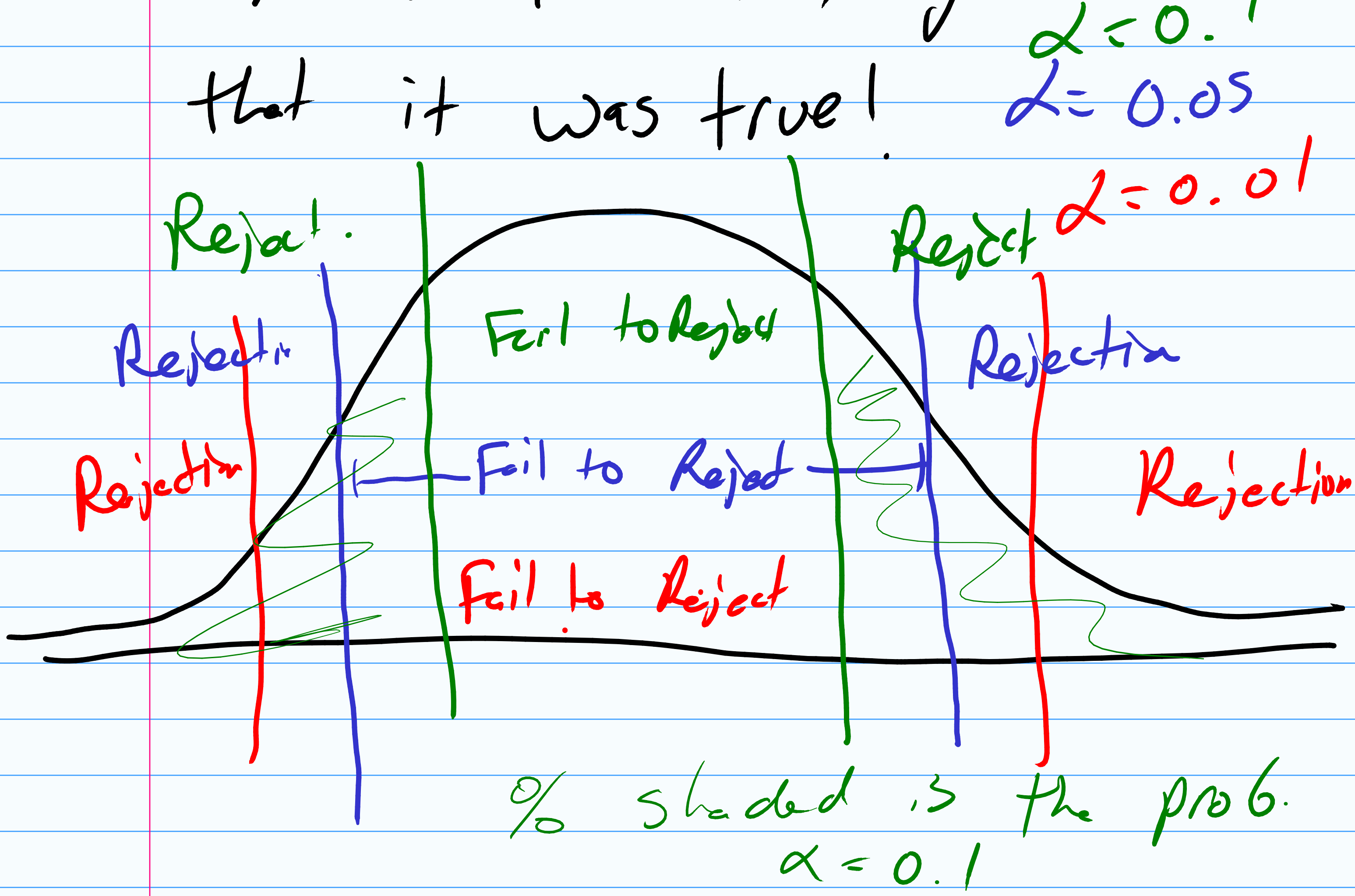


If T_0^+ is outside the bounds we then say that the test statistic (T_0^+) is unlikely to come from the null distribution

1. If it is w/i the dist. we Fail to Reject H_0!

2. If it is outside the dist. we then Reject H_0!

Given the 2 scenarios, α controls the probability of, rejecting the H_0, given that it was true!



This method is know as the critical value approach.

P-Value Approach.

$H_0: P_{non} - P_{smoke} = 0$

$H_a: P_{non} - P_{smoke} \neq 0$

① Define a significance level.

$\alpha = 0.05$

② Compute a test statistic.

T_0^{t}

$H_a: p_1 - p_2 \neq 0$

③ Compute the P-value

$H_a: p_1 - p_2 \neq 0$

$P(T_0 > |T_0^{t}|) =$ p-value.

$$P(T_0 > +|T_0^+|) + P(T_0 < -|T_0^+|)$$

If symmetrical

$$\text{P-value} = 2P(T_0 > |T_0^+|)$$

Reject H_0	Fail to Reject H_0
$p < \alpha$	$p \geq \alpha$

P-value

$H_a: p_1 - p_2 < 0$

Test statistic

$$P(T_0 < T_0^+) = \int_{-\infty}^{T_0^+} f(T_0)\, dT_0$$

$H_a: p_1 - p_2 > 0$

Test statistic

$$P(T_0 > T_0^+) = \int_{T_0^+}^{\infty} f(T_0)\, dT_0$$

Test a population mean μ

Assumptions.

①. $X = \{X_1, \dots, X_n\} \overset{iid}{\sim} N(\mu_0, \sigma^2)$

σ^2 is known

We are interested in knowing if the data points came from our theorized μ_0 value.

①

$H_0: \mu = \mu_0$

$H_a: \mu \neq \mu_0$

$\mu > \mu_0$

$\mu < \mu_0$

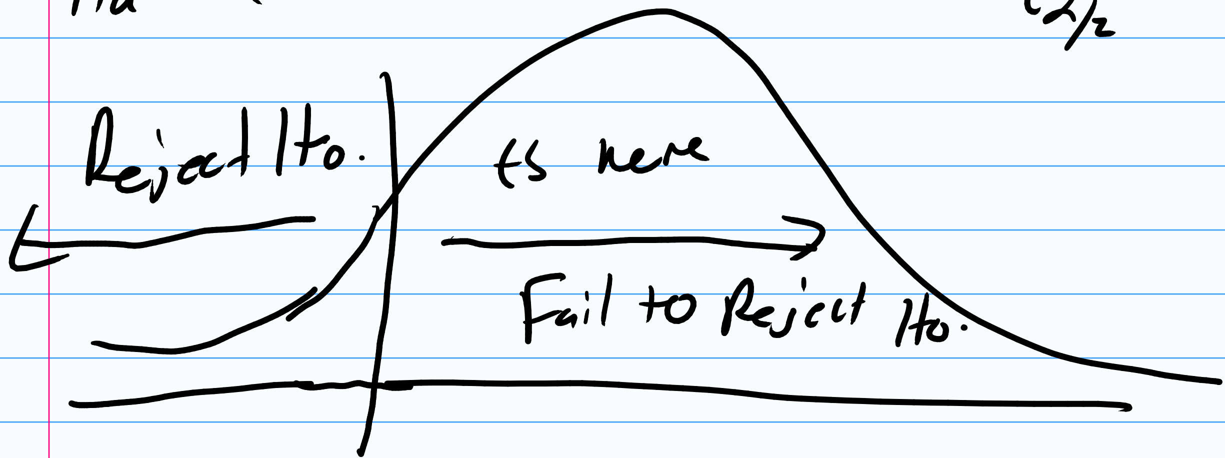
②. construct a test statistic that is appropriate to test the null hypothesis.

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

$$\bar{X} - \mu_0 \sim N(0, \sigma^2/n)$$

$$ts = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}} \sim N(0, 1)$$

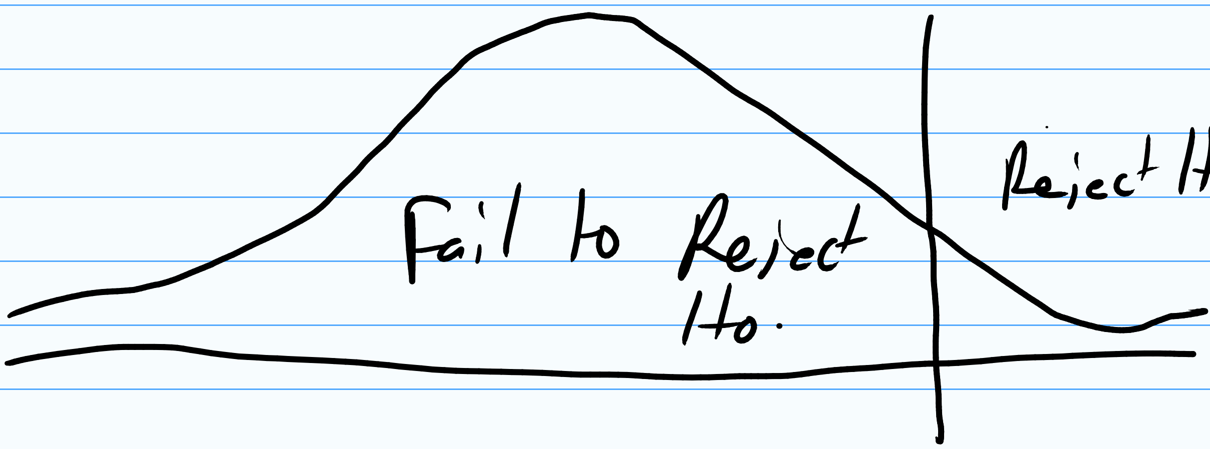
H_a $\mu < \mu_0$ $CV = z_{\alpha/2}$



$z_{\alpha/2}$

$$P(Z < ts) = p\text{-value}$$

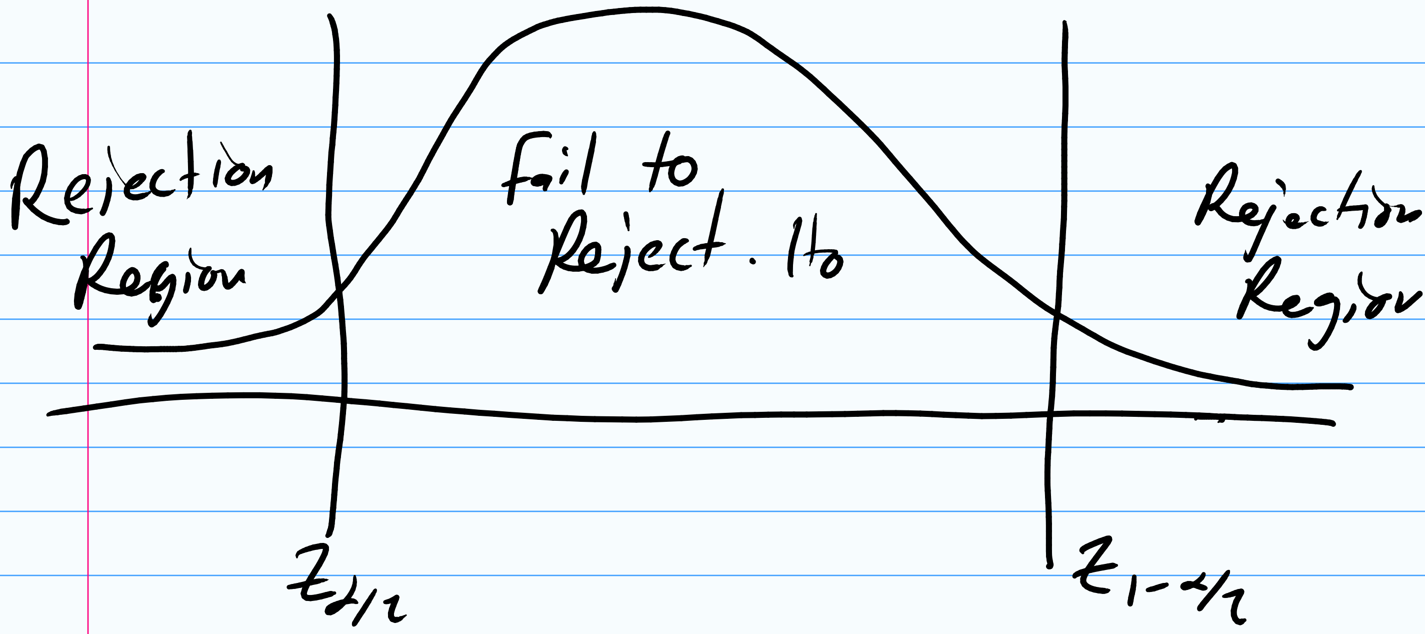
H_a: $\mu > \mu_0$ $z_{1-\alpha/2}$ CV



$z_{1-\alpha/2}$

$$P(Z > ts)$$

$H_a \; \mu \neq \mu_0$ $\qquad z_{\alpha/2}, \; z_{1-\alpha/2}$



$$\text{p-value} = 2 \cdot P(Z > |ts|)$$