Recall :-

- X - population of interest ; assume that X has either
  p.m.f $f(\cdot \mid p)$ / p.d.f $f(\cdot \mid p)$ $\}$ $p \in \mathcal{P} \subseteq \mathbb{R}^d$.

- Conjecture :- Is $p \in \mathcal{P}_0 \subseteq \mathcal{P}$?
  [Null hypothesis]

Necessity to find best possible

- Test to verify the conjecture :-
  - Sample $X_1, X_2, \ldots, X_n$ i.i.d. X
  - Develops a "test statistic" ≡ function of $X_1, \ldots, X_n$ ≡ Distribution will depend on $p \in \mathcal{P}$.

Distribution is fully known when $p \in \mathcal{P}_0$

  - Verify :- one tries to observe if indeed the statistic could have arisen from $p \in \mathcal{P}_0$

  - Quantify the degree of possibility through a probability called the p-value.

- Null hypothesis : Conjecture one makes about the population

- Alternate hypothesis :- this will specify a manner in which the null is erroneous

- p-value :- probability that sample would be atleast as from expectation as was actually observed.
  - small p-value = indicates sample is highly unusual $\Rightarrow$ Null is perhaps not true.
  - large p-value ... consistent with null hypothesis

Z-test :- Test for sample mean when σ is known.

$X \sim \text{Normal}(\mu, \sigma^2)$

I Null hypothesis : $\mu = c$

Alternate hypothesis: $\mu > c$

let $X_1, \ldots X_n$ be i.i.d samples from X

Compute :- $\sqrt{n}\left(\frac{\overline{X} - c}{\sigma}\right)$ $\quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$

Fix level of significance α : $\alpha = 0.05$

(threshold for p-value)

$Z \sim \text{Normal}(0,1)$

- $\mathbb{P}\left(Z \geq \sqrt{n}\left(\frac{\overline{X} - c}{\sigma}\right)\right) < \alpha$

reject the null hypothesis.

(II)

Null hypothesis : $\mu = c$

Alternate hypothesis : $\mu \neq c$.

Sample $X_1, \ldots, X_n$ from the population

let $Y_1, \ldots, Y_n$ be iid X

$$\Rightarrow \quad Z = \sqrt{n}\left(\frac{\overline{Y} - c}{\sigma}\right) \sim \text{Normal}(0,1)$$

Compare : how far is $\overline{X}$ from c?

i.e need to understand $|\overline{X} - c|$

$$\mathbb{P}\left(|\bar{Y} - c| \geq |\bar{X} - c| \right)$$

$$= \mathbb{P}\left(\sqrt{n} \frac{(|\bar{Y} - c|)}{\sigma} \geq \sqrt{n}\left(\frac{|\bar{X} - c|}{\sigma}\right) \right)$$

$$= \mathbb{P}\left(|Z| \geq \sqrt{n}\left(\frac{|\bar{X} - c|}{\sigma}\right) \right)$$

Fix $\alpha \in (0,1)$

step 1:- Sample $X_1 \ldots X_n$ & $\frac{\sqrt{n}(\bar{X} - c)}{\sigma}$ = Compute

step 2:- Compute $\mathbb{P}\left(|Z| \geq \left| \sqrt{n} \frac{(\bar{X} - c)}{\sigma} \right| \right)$

step 3:- if "$\mathbb{P}\left(|Z| \geq \left| \sqrt{n} \frac{(\bar{X} - c)}{\sigma} \right| \right) < \alpha$

then reject the null hypothesis.

(III) Null hypothesis: $\mu = c$

Alternate hypothesis: $\mu < c$

Ex: Device appropriate test.

t-test :- for sample mean when σ is unknown

$X \sim$ Normally distributed with mean μ, variance σ^2

μ and σ are unknown.

Null hypothesis :- $\mu = c$

Alternate hypothesis : $\mu \geq c$

let $Y_1, \ldots, Y_n$ be i.i.d X "mimic" sample

under Null : $Y_i \sim \text{Normal}(c, \sigma^2)$

Known :- $T := \sqrt{n}\,\dfrac{(\bar{Y} - c)}{S} \sim t_{n-1}$

ution of test statistic
on under null
thesis

$\bar{Y} \geq \bar{X}$)

$$P\left(T \geq \sqrt{n}\,\frac{(\bar{X} - c)}{S} \right)$$

Fix $\alpha \in (0,1)$

step 1:- Sample $X_1, \ldots X_n$ & $\dfrac{\sqrt{n}(\bar{X}-c)}{S}$ = Compute

step 2:- Compute $\mathbb{P}\left(T \geq \dfrac{\sqrt{n}(\bar{X}-c)}{S}\right)$ where $T \sim t_{n-1}$ distribution

step 3:- if $\mathbb{P}\left(T \geq \dfrac{\sqrt{n}(\bar{X}-c)}{S}\right) < \alpha$ then reject the null hypothesis.

Ex:- $X \sim$ Normally distributed with variance σ

μ and σ are unknown.

(II) Null hypothesis :- $\mu = c$

Alternate hypothesis : $\mu < c$

(III) Null hypothesis :- $\mu = c$

Alternate hypothesis : $\mu \neq c$

Question:

Device test as above ?

χ^2-test : Test for sample variance

$X \sim$ Normal (μ, σ^2) with σ – unknown.

$X_1 \ldots, X_n$ be sample from the population

$Y_1 \ldots Y_n$ be iid X.

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

(Sample variance) $$S_Y^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} (Y_i - \bar{Y})^2 \right)$$

$E[S_Y^2] = \sigma^2$. (seen before)

Distribution of S_Y^2 :

χ^2_n – [chi-square with n degrees of freedom] – is a random

with p.d.f given by

$$f(x) = \frac{2^{-n/2}}{\Gamma(n/2)} x^{n/2-1} e^{-x/2}$$

Before:

$n=1$, $X \sim$ Normal$(0,1)$

$X^2 \sim \chi^2_1$

[Theorem 8.5.9 in Book]

$\frac{(n-1) S_y^2}{\sigma^2} \sim$ has χ^2_{n-1} distribution.

[under the null hypothesis one can work with the test - statistic]

Null hypothesis :- $\sigma = c$

Alternative hypothesis :- $\sigma > c$

$S_x^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

$$\mathbb{P}(S_y^2 \geq S_x^2) = \mathbb{P}\left(\frac{n-1}{c^2} S_y^2 \geq \frac{(n-1) S_x^2}{c^2} \right)$$

$$= \mathbb{P}\left(\chi^2_{n-1} \geq \frac{(n-1) S_x^2}{c^2} \right)$$

step 1 :- Fix $\alpha \in (0,1) \equiv$ level of significance

$X_1, \ldots, X_n$ from the population

Compute $\frac{(n-1) S_x^2}{c^2}$

Step 2 :- Compute $\mathbb{P}\left(\chi^2_{n-1} \geqslant \frac{(n-1) S_x^2}{c^2} \right)$

Step 3 :- If $\mathbb{P}\left(\chi^2_{n-1} \geqslant \frac{(n-1) S_x^2}{c^2} \right) < \alpha$

then reject the null hypothesis.

Ex :- X ~ Normally distributed with mean μ variance σ

μ and σ are unknown.

II Null hypothesis :- $\sigma = c$

alternate hypothesis : $\sigma < c$

III Null hypothesis :- $\sigma = c$

alternate hypothesis : $\sigma \neq c$

Question : Device test as above.