Chapter - 6

12th March '2024

Normal Theory :→

In this chapter, we look at Normal random variables under different conditions. We want to estimate parameters, build models, and perform hypothesis tests on μ and σ^2 in these differing conditions.

Section 6.1: Basic Assumptions.

1) Let, $X_1, X_2, \ldots, X_n$ be independent random variables with $X_i \sim N(\mu_i, \sigma_i^2)$ and, let $a_1, a_2, \ldots, a_n$ be constants. Then,

$$\sum_{i=1}^{n} a_i X_i \sim N\Big(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\Big)$$

$Var(A+B) = Var(A) + Var(B) + \cancel{Cov(A,B)}$

$[Cov(A,B) = 0, \text{ when } A \perp B]$

2) Let, $Z_1, Z_2, \ldots, Z_n \overset{iid}{\sim} N(0,1)$. Then

$$\sum_{i=1}^{n} Z_i^2 \sim \chi^2_{(n)}, \text{ because } Z_i^2 \sim \chi^2_{(1)}, \forall i$$

Let's assume we have "n" independent observations $y_1, y_2, \ldots, y_n$. Each of these is a realization of a R.V $Y_1, Y_2, \ldots, Y_n$, where we assume $Y_i \sim N(\mu_i, \sigma^2)$

Constant across the measurements.

Constant, but each R.V has its own mean.

Using (1), we can express $Y_i = \mu_i + \varepsilon_i$ Called "Error" or "residual"

Constant

Random, $\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$

$$E(Y_i) = E(\mu_i + \varepsilon_i)$$
$$= E(\mu_i) + E(\varepsilon_i)$$
$$= \mu_i + 0 = \mu_i$$

$$Var(Y_i) = Var(\mu_i + \varepsilon_i)$$
$$= Var(\mu_i) + Var(\varepsilon_i) + Cov(\mu_i, \varepsilon_i)$$
$$= 0 + \sigma^2 + 0 = \sigma^2$$

- How many parameters do we need to estimate?

  ↳ $\mu_1, \mu_2, \ldots, \mu_n, \sigma^2$

  So, $k = \begin{cases} n & ; \text{if } \sigma \text{ is known} \\ n+1 & ; \text{if } \sigma \text{ is unknown} \end{cases}$

- We will reduce the # of parameters with various models.

⟺ One sample model : n measurements taken under the same conditions.

Assume, $\mu_1 = \mu_2 = \cdots = \mu_n = \mu$ unknown

in Lecture notes, "α"

Now, $\boxed{q = 1}$ unknown parameter to estimate, assuming σ^2 known.

⟺ Two sample model : Two groups of sample measurements.

y_{1i}, $i = 1, 2, \cdots, n_1$; n_1 = sample size of group 1.

y_{2j}, $j = 1, 2, \cdots, n_2$; n_2 = sample size of group 2.

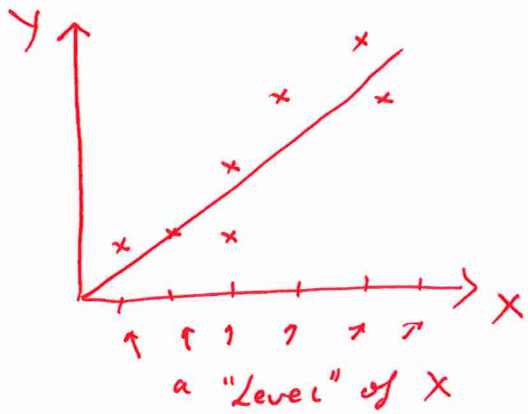
Assume: $\left.\begin{aligned} \mu_{11} = \mu_{12} = \cdots\cdots = \mu_{1n_1} = \mu_1 \\ \mu_{21} = \mu_{22} = \cdots\cdots = \mu_{2n_2} = \mu_2 \end{aligned}\right\}$ unknown

So, $q = 2$ unknown parameters, assuming σ_1^2 and σ^2 are known.

μ_1, μ_2 : in some contexts; it makes more sense to say, $\mu_1 = \alpha$, $\mu_2 = \alpha + \beta$, where, α, β unknown.

Straight Line Model :- "n" measurements across varying conditions.

Assume: $\mu_i = \alpha + \beta X_i$, where X_i is your "predictor variable"



Assuming, σ^2 is known.

$$\boxed{q = 2}$$

See Co-op files in Supplementary Materials (ch-6).