Lecture 27

niceguy

March 20, 2023

1 Paired Observations

With 2 samples of the same size n, each pair has means from each.

Example 1.1. The blood pressure of participants before and after a medical trial.

We define the difference $D_i = X_i - Y_i$. Now

$$\mathrm{var}(D_i) = \mathrm{var}(X_i - Y_i) = \sigma_X^2 + \sigma_Y^2 - 2\mathrm{cov}(X_i, Y_i)$$

A lower variance leads to a narrower confidence interval. We have

$$Z = \frac{\overline{D} - (\mu_X - \mu_Y)}{\sqrt{\frac{1}{n}(\sigma_X^2 + \sigma_Y^2 - 2\mathrm{cov}(X, Y))}}$$

2 Bernoulli Random Variable

It is a binomial with

$$P(Y_i = 1) = p \forall i \in \mathbb{Z}^+$$

(think coin flips). We define

$$X = \sum_{i=1}^{n} Y_i$$

and

$$b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$$

We have

$$\mu = np, \sigma^2 = np(1-p)$$

Let

$$Z = \frac{X - np}{\sqrt{np(1-p)}}$$

As $n \to \infty$, the distribution of Z goes towards $n(z;0,1)$ by the Central Limit Theorem.

Now we estimate $\hat{p} = \frac{X}{n}$. It is unbiased, as the mean is

$$\mu = E\left[\frac{X}{n}\right] = \frac{1}{n}E[X] = \frac{1}{n}np = p$$

and

$$\sigma^2 = \frac{1}{n^2}\sigma_X^2 = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

Now the statistic

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

tends to the normal distribution as $n \to \infty$.

The confidence interval becomes

$$\begin{aligned} 1 - \alpha &= p(-z_{\frac{\alpha}{2}} \leq Z \leq z_{\frac{\alpha}{2}}) \\ &= p(-z_{\frac{\alpha}{2}} \leq \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \leq z_{\frac{\alpha}{2}}) \end{aligned}$$

where

$$z_{\frac{\alpha}{2}} = -\Phi^{-1}\left(\frac{\alpha}{2}\right)$$

Rearranging,

$$1 - \alpha = P\left(\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) \leq P\left(p \leq \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

We want this confidence intervala to be no wider than 2δ. Rearranging gives

$$n \geq \frac{z_{\frac{\alpha}{2}}^2}{4\delta^2}$$

where we use the fact that

$$\hat{p}(1-\hat{p}) \leq \frac{1}{4}$$

because $\hat{p}$ is positive.