Conditional Power

Chao Cheng

September 5, 2023

Contents

1 Introduction

The conditional power is defined as the probability to reject the null hypothersis at **Final** analysis given the current collected data.

2 Mean Test

2.1 Normal distribution, variance known

2.1.1 One-sample mean test

Let the sample $x_1, \cdots, x_n \sim N(\mu, \sigma^2)$, where σ^2 is known. To simplify the discussion here, we assume an one-sided test

$$H_0: \ \mu \le \mu_0, \quad H_1: \ \mu > \mu_0$$

The reject rule is rejecting H_0 if $\bar{x} > c$ for some c. Since $\bar{x} \sim N(\mu, \sigma^2/n)$, c can be determined by type-I error α:

$$P(\bar{x} > c | H_0) \le \alpha.$$

By Neyman-Pearson Lemma, and stochastical monotone property of normal distribution, to fully utilize the type-I error, it is

$$P(\bar{x} > c | \mu = \mu_0) = P\left(\frac{\bar{x} - \mu_0}{\sqrt{\sigma^2/n}} > \frac{c - \mu_0}{\sqrt{\sigma^2/n}}\right) = P\left(Z > \frac{c - \mu_0}{\sqrt{\sigma^2/n}}\right) = \alpha.$$

Hence $c = \mu_0 - z_\alpha \sqrt{\sigma^2/n}$, where $z_\alpha = \Phi^{-1}(\alpha)$ is the (left-)α quantile of normal distribution. For more details about how this test is constructed, one can refer to the **ttest.pdf** notes in this repo.

The probability to reject the null hypothersis at some given μ_1 can be computed as

$$\begin{aligned}
&P(\bar{x} > c|\mu = \mu_1) \\
=&P\left(\frac{\bar{x} - \mu_1}{\sqrt{\sigma^2/n}} > \frac{c - \mu_1}{\sqrt{\sigma^2/n}}\right) \\
=&P\left(Z > -z_\alpha + \frac{\mu_0 - \mu_1}{\sqrt{\sigma^2/n}}\right) \\
=&1 - \Phi\left(-z_\alpha + \frac{\mu_0 - \mu_1}{\sqrt{\sigma^2/n}}\right) = 1 - \beta,
\end{aligned}$$

where β is the type-II error. Then the samplesize is determined by

$$n = \frac{\sigma^2 (z_\alpha + z_\beta)^2}{(\mu_0 - \mu_1)^2}.$$

Now, assume we have collected n_1 samples, which means there are $n_2 = n - n_1$ more samples to come. And the current statistics is $\bar{x}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i$ and the final statistics is $\bar{x} = \frac{1}{n_1+n_2}\sum_{i=1}^{n_1+n_2} x_i$. The statistics from the second part is $\bar{x}_2 = \frac{1}{n_2}\sum_{i=n_1+1}^{n_1+n_2} x_i$. Note that these ($\bar{x}_1$, $\bar{x}_2$ and $\bar{x}$) are all sufficient statistics for μ (based on part1, part1 and all parts).

Here we know

$$\bar{x}_1 \sim N\left(\mu, \sigma^2/n_1\right), \quad \bar{x}_2 \sim N\left(\mu, \sigma^2/n_2\right), \quad \bar{x}_1 \perp \bar{x}_2.$$

And $\bar{x} = \frac{n_1}{n_1+n_2}\bar{x}_1 + \frac{n_2}{n_1+n_2}\bar{x}_2$. So the conditional power, which is the probability to reject the hypothesis at final analysis given the current collected data is

$$\begin{aligned}
&P(\bar{x} > c|\bar{x}_1) \\
=&P\left(\frac{n_1}{n_1+n_2}\bar{x}_1 + \frac{n_2}{n_1+n_2}\bar{x}_2 > c \,\middle|\, \bar{x}_1\right) \\
=&P\left(\bar{x}_2 > \frac{(n_1+n_2)c - n_1\bar{x}_1}{n_2}\right) \\
=&P\left(\frac{\bar{x}_2 - \mu}{\sqrt{\sigma^2/n_2}} > \frac{\frac{(n_1+n_2)c - n_1\bar{x}_1}{n_2} - \mu}{\sqrt{\sigma^2/n_2}}\right) \\
=&P\left(Z > \frac{-z_\alpha\sqrt{(n_1+n_2)\sigma^2} + n_1(\mu_0 - \bar{x}_1) + n_2(\mu_0 - \mu)}{\sqrt{n_2\sigma^2}}\right) \\
=&1 - \Phi\left(\frac{-z_\alpha\sqrt{(n_1+n_2)\sigma^2} + n_1(\mu_0 - \bar{x}_1) + n_2(\mu_0 - \mu)}{\sqrt{n_2\sigma^2}}\right)
\end{aligned} \tag{1}$$

Note: in previous discussions there are some abuse of the notation. $\bar{x}_1$ and $\bar{x}_2$ can both refer to the random variable, and there realization. But I hope the context is clear and not much confusion.

2.1.2 Two-sample mean test

3 Rate Test

4 Log Rank Test

5 Sample Size Re-estimation/Re-calculation