11. P-value Exercises

A Formula for the p-value

1/1 point (graded)

Consider a statistical experiment $X_1, \ldots, X_n \sim N(\mu, 1)$, where μ is an unknown parameter. We will hypothesis test on the parameter μ by setting $H_0 : \mu = 0$ and $H_1 : \mu \neq 0$. Our test is designed to be

$$\psi_n = \mathbf{1}\left(|\sqrt{n}\overline{X}_n| \geq q_{\eta/2}\right)$$

where q_η denotes the $1-\eta$ quantile of a standard Gaussian.

Let $\Phi(x)$ denote the CDF of the standard Gaussian distribution $N(0,1)$.

Which of the following is the p-value for this experiment?

- $\Phi\left(|\sqrt{n}\overline{X}_n|\right)$
- $2\Phi\left(|\sqrt{n}\overline{X}_n|\right)$
- $1-\Phi\left(|\sqrt{n}\overline{X}_n|\right)$
- $2(1-\Phi\left(|\sqrt{n}\overline{X}_n|\right)$ ✔

Solution:

For notational convenience, let

$$T_n = \sqrt{n}\overline{X}_n.$$

As in the previous examples, the p-value is obtained by setting $q_{\eta/2} = \sqrt{n}\overline{X}_n$ and solving for η. Since, by definition, $\eta = P(Z \geq q_\eta)$ for $Z \sim N(0,1)$, we have that the p-value is given by

$$\eta = 2P\left(Z > q_{\eta/2}\right) = 2P\left(Z > |T_n|\right) = 2\left(1 - \Phi\left(|T_n|\right)\right).$$

The last choice is correct.

Remark: In general, we will compute p-values using the CDF of the underlying distribution. Although the p-value has a rather complicated definition, in the case of most of the models we work with, it can be computed in a relatively straightforward fashion.

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

Concept Check: Properties of p-values

0/1 point (graded)
Which of the following are true statements regarding p-values? (Choose all that apply.)

- [ ] The p-value represents a **tipping point** in the sense that for any level smaller than the p-value, our test would fail to reject the null hypothesis based on the data. ✔ fix p-value, change alpha level
- [x] The smaller the p-value, the more confidently one can reject the null hypothesis. ✔
- [x] The p-value is computed based on the sample that we observe. ✔
- [x] One way that scientists and companies have, in some instances, artifically lowered p-values is by specifying the null and alternative hypotheses *after* observing the data. ✔

✖

Solution:

All of the above choices are correct. We examine each individually.

- The first choice elaborates on the definition of the (asymptotic) p-value, which is the smallest (asymptotic) level at which a test ψ will reject the null hypothesis. Hence, for any asymptotic level below the p-value, our test will **fail to reject** on our observed sample. This is a slight restatement of the 'golden rule' described in the slides.

- The second choice is correct, "The smaller the p-value, the more confidently one can reject the null hypothesis." Suppose, for the sake of example, that

$$\psi = \mathbf{1}\left(|T_n| > q_{\eta/2}\right),$$

$T_n \xrightarrow[n\to\infty]{(d)} N(0,1)$, and q_η represents the $1-\eta$ quantile of a standard Gaussian.

Then the asymptotic p-value is the smallest value we can plug in for η such that

$$|T_n| \geq q_{\eta/2}.$$

If η is the smallest such value, this implies that $T_n = q_{\eta/2}$. Now solving for η, if it turns out to be very small, this means $q_{\eta/2}$ is very large and so T_n lies deeper in the tails of the distribution $N(0,1)$. Intuitively, we should think of the tails of $N(0,1)$ as housing the 'rare events', so by this reasoning, it makes sense to interpret a small p-value as an indicator that we observed a rare event under H_0. Hence, we may be convinced to reject the null under such an observation.

- "The p-value is computed based on the sample that we observe." is correct. It is important to keep in mind how it is computed, and the previous problem serves as a good example of the typical strategy.

- It is very important in practice that one specifies the null and alternative hypotheses *before* conducting the experiment. Otherwise, it is possible to 'tweak' the hypotheses. This can artifically result in a lower p-value, which would favor the scientist or company's desired conclusion. For example, if

  - H_0: a drug is no more effective than placebo
  - H_1: a drug is more effective than placebo

  then a drug company would consider rejecting H_0 as a success, since it would validate their product. Since a lower p-value favors rejection, artifically lowering them is a fraudulent method of attaining a favorable outcome.

提交 你已经尝试了2次（总共可以尝试2次）

ℹ Answers are displayed within the problem

讨论

显示讨论