Master of Data Science Online Programme
Course: Basic Statistics
SGA #4: Hypothesis testing meets confidence intervals

Student: Andrei Batyrov (Fall 2022)
arbatyrov@edu.hse.ru

Higher School of Economics (HSE)
National Research University
Faculty of Computer Science

June 11, 2023

Contents

Problem 1

There is a connection between confidence intervals and hypothesis testing. Assume that we have an i.i.d. sample $x = (x_1, \dots, x_n)$ from some random variable X with finite variance. Consider one-sample t-test with null hypothesis $\mathbb{E}X = \mu_0$ and symmetric alternative. For simplicity, let us assume that n is large enough and replace T-distribution with standard normal distribution. Assume that one found confidence interval $\mathcal{I}$ for $\mathbb{E}X$ with confidence level 95%. Prove that standard decision-making procedure of t-test is equivalent to the following: reject null hypothesis if and only if μ_0 does not belong to $\mathcal{I}$. Follow the plan:

1. Assume that null hypothesis holds. We believe that t-statistics in this case is distributed according to standard normal law (due to assumption that n is large). Recall that t-statistics for sample x is defined as
$$t \approx \frac{\bar{x} - \mu_0}{SD(x)}\sqrt{n}. \tag{1}$$
2. If μ_0 does not lie in $\mathcal{I}$, either μ_0 is larger than the right endpoint of $\mathcal{I}$ or μ_0 is smaller than the left endpoint of $\mathcal{I}$. Let us consider the latter case.
3. Consider event "μ_0 is smaller than the left endpoint of $\mathcal{I}$". Write this condition as an inequality using $\mu_0, \bar{x}, SD(x), n$ and a constant 1.96. (Recall that we assume that null hypothesis holds.)
4. Transform this inequality such that it becomes $(\dots) > 1.96$. Does the left-hand part look similar to something?
5. Recall why we use number 1.96, how it is connected to standard normal distribution.
6. Find probability that μ_0 is smaller than the left endpoint of $\mathcal{I}$ provided that null hypothesis holds.
7. Find probability that μ_0 does not lie in $\mathcal{I}$ provided that null hypothesis holds.
8. Assume we are following rule "reject null hypothesis if and only if μ_0 does not belong to $\mathcal{I}$." Find probability that we falsely reject null hypothesis provided that it is true.
9. Explain in what cases (in terms of $\bar{x}$) will we reject null hypothesis if we follow mentioned rule.
10. Explain that this rule is equivalent to the rule used in ordinary two-sided one-sample t-test.

Solution

1. Let's consider our hypotheses about μ_0:

   $\mathcal{H}_0 : \mathbb{E}X = \mu_0$, population mean is equal to μ_0,

   $\mathcal{H}_1 : \mathbb{E}X \neq \mu_0$, population mean is not equal to μ_0, i.e., symmetric two-sided alternative.
2. As per [1], the confidence interval the population mean μ_0 belongs to is
$$\mathcal{I}(x) = (\bar{x} - s, \bar{x} + s), \tag{2}$$
   where s is a parameter such that the probability that μ_0 belongs to $\mathcal{I}$
$$P(\mu_0 \in (\bar{x} - s, \bar{x} + s)) \tag{3}$$
   is equal to some agreed value, called confidence level $\in [0, 1]$.
3. For the confidence level 0.95 (95%), as per [1],
$$s = 1.96 \times \frac{SD(x)}{\sqrt{n}}, \tag{4}$$
   where $SD(x)$ is the unbiased standard deviation of the sample x and n is the sample size of x. Let's consider event "μ_0 is smaller than the left endpoint of $\mathcal{I}$". As per (2) and (4), it can be expressed as
$$\mu_0 < \bar{x} - s \Rightarrow \mu_0 < \bar{x} - 1.96 \times \frac{SD(x)}{\sqrt{n}}. \tag{5}$$

4. Rearranging (5) yields

$$\begin{aligned} \mu_0 &< \bar{x} - 1.96 \times \frac{SD(x)}{\sqrt{n}} \\ 1.96 \times \frac{SD(x)}{\sqrt{n}} &< \bar{x} - \mu_0 \\ \bar{x} - \mu_0 &> 1.96 \times \frac{SD(x)}{\sqrt{n}} \\ \frac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} &> 1.96. \end{aligned} \tag{6}$$

   The left-hand side of (6) is exactly the t-score (1)!

5. The "magic" value 1.96 is in fact simply the value of z-score of the standard normal distribution, such that

$$P(z > 1.96) = 1 - P(z \leq 1.96) = 1 - CDF_{\mathcal{N}(0,1)}(1.96) = 1 - 0.975 = 0.025. \tag{7}$$

   In other words, 1.96 is the 97.5th percentile of $Z \sim \mathcal{N}(0, 1)$. Since, the standard normal distribution is symmetrical around μ_0,

$$P(z < -1.96) = CDF_{\mathcal{N}(0,1)}(-1.96) = 0.025. \tag{8}$$

   In other words, (-1.96) is the 2.5th percentile of $Z \sim \mathcal{N}(0, 1)$. Strictly speaking, we should use the approximately equals sign here, because ± 1.96 is the rounded value of exact 2.5th and 97.5th percentiles. However, since ± 1.96 is very close to true values, for simplicity, we are using the equal sign here and below.

6. As per (5), (6), (1), and (7), and recalling that we assumed that $T \sim \mathcal{N}(0, 1)$, the probability of the event "μ_0 is smaller than the left endpoint of $\mathcal{I}$ provided that null hypothesis holds" is

$$P(\mu_0 < \bar{x} - s \,|\, \mathcal{H}_0) = P\left(\frac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} > 1.96\right) = P(t > 1.96) \approx P(z > 1.96) = 0.025. \tag{9}$$

7. The probability of the event "μ_0 does not lie in $\mathcal{I}$ provided that null hypothesis holds" is

$$P(\mu_0 < \bar{x} - s \cup \mu_0 > \bar{x} + s \,|\, \mathcal{H}_0) = P(\mu_0 < \bar{x} - s) + P(\mu_0 > \bar{x} + s). \tag{10}$$

   To find $P(\mu_0 > \bar{x} + s)$, let's do rearrangements similar to (6):

$$\begin{aligned} \mu_0 &> \bar{x} + 1.96 \times \frac{SD(x)}{\sqrt{n}} \\ -1.96 \times \frac{SD(x)}{\sqrt{n}} &> \bar{x} - \mu_0 \\ \bar{x} - \mu_0 &< -1.96 \times \frac{SD(x)}{\sqrt{n}} \\ \frac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} &< -1.96. \end{aligned} \tag{11}$$

   Finally, As per (5), (6), (11), (1), (7), (8), and recalling that we assumed that $T \sim \mathcal{N}(0, 1)$, (10) yields:

$$\begin{aligned} P(\mu_0 < \bar{x} - s \cup \mu_0 > \bar{x} + s \,|\, \mathcal{H}_0) &= P(\mu_0 < \bar{x} - s) + P(\mu_0 > \bar{x} + s) \\ &= P\left(\frac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} > 1.96\right) + P\left(\frac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} < -1.96\right) \\ &= P(t > 1.96) + P(t < -1.96) \\ &\approx P(z > 1.96) + P(z < -1.96) \\ &= 0.025 + 0.025 = 0.05. \end{aligned} \tag{12}$$

8. As per [5], probability that we falsely reject null hypothesis provided that it is true is the probability of Type I error. For two-sided alternative hypothesis this probability is

$$P(\text{reject } \mathcal{H}_0 \,|\, \mathcal{H}_0 \text{ holds}) = P(t > t_{crit} \,|\, \mathcal{H}_0) + P(t < -t_{crit} \,|\, \mathcal{H}_0) = \alpha, \tag{13}$$

where t_{crit} is some value such that (13) is true for some fixed value called significance level α. Recall, that as per (12), the probability of the event "μ_0 does not lie in $\mathcal{I}$ provided that null hypothesis holds" is $P(\mu_0 < \bar{x} - s \cup \mu_0 > \bar{x} + s \,|\, \mathcal{H}_0) \approx P(z > 1.96) + P(z < -1.96)$, i.e we used $t_{crit} \approx z_{crit} = 1.96$ as the critical value to determine if μ_0 does not belong to $\mathcal{I}$: $t_{crit} \approx z_{crit} = 1.96$ for "μ_0 is smaller than the left endpoint of $\mathcal{I}$" and $-t_{crit} \approx -z_{crit} = -1.96$ for "μ_0 is larger than the right endpoint of $\mathcal{I}$". Thus, the probability of Type I error under the rule "reject null hypothesis if and only if μ_0 does not belong to $\mathcal{I}$" is

$$\begin{aligned} P(t > t_{crit} \,|\, \mathcal{H}_0) + P(t < -t_{crit} \,|\, \mathcal{H}_0) &= P(t > 1.96) + P(t < -1.96) \\ &\approx P(z > 1.96) + P(z < -1.96) \\ &= 0.025 + 0.025 = 0.05. \end{aligned} \tag{14}$$

Note, that the results (12) and (14) are essentially the same, i.e., the probability that μ_0 does not lie in $\mathcal{I}$ is equal to the probability of falsely rejecting the null hypothesis provided that it is true (probability of Type I error). In fact, we have "reverse-engineered" α, i.e., if we had set $\alpha = 0.05$, we would get $t_{crit} \approx z_{crit} = 1.96$ for $\alpha/2$.

9. To reject the null hypothesis we need to compare the so-called p-value [2] with the significance level α: if p-value $< \alpha$, we reject $\mathcal{H}_0$ in favor of $\mathcal{H}_1$, and do not reject otherwise. To satisfy this condition for μ_0 being outside of $\mathcal{I}$, we have to be on either side of the critical region of t, i.e., $p\text{-value} = P\left(t = \dfrac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} > 1.96\right) + P\left(t = \dfrac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} < -1.96\right) < \alpha \Rightarrow$ since CDF is a monotonically increasing function [4] $\Rightarrow t = \dfrac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} > 1.96$ or $t = \dfrac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} < -1.96$. Rearranging these expressions for $\bar{x}$ yields:

$$\bar{x} > \mu_0 + 1.96 \times \frac{SD(x)}{\sqrt{n}} \text{ or } \bar{x} < \mu_0 - 1.96 \times \frac{SD(x)}{\sqrt{n}}. \tag{15}$$

We will reject the null hypothesis, if either of the conditions (15) is satisfied.

10. Recalling that for passing the two-tailed one-sample t-test [3] we need to guarantee that p-value $< \alpha$ and for that we need to be on either side of the critical region of t, i.e. $p\text{-value} = P\left(t = \dfrac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} > t_{crit}\right) + P\left(t = \dfrac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} < -t_{crit}\right) < \alpha \Rightarrow$ since CDF is monotonically increasing function [4] $\Rightarrow t = \dfrac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} > t_{crit}$ or $t = \dfrac{\bar{x} - \mu_0}{SD(x)}\sqrt{n} < -t_{crit}$, where t_{crit} is determined by the fixed significance level α divided by 2. We can easily see that the condition for rejecting $\mathcal{H}_0$ is essentially the same as (15):

$$\bar{x} > \mu_0 + t_{crit} \times \frac{SD(x)}{\sqrt{n}} \text{ or } \bar{x} < \mu_0 - t_{crit} \times \frac{SD(x)}{\sqrt{n}}. \tag{16}$$

Answer

We have shown by (12) and (14), that the probability of falsely rejecting the null hypothesis provided that it is true (probability of Type I error) is essentially the same as the probability that μ_0 does not lie in $\mathcal{I}$. Thus, in order to pass the two-tailed one-sample t-test we reject the null hypothesis (controlling Type I error) if and only if μ_0 does not belong to $\mathcal{I}$, which is equivalent to checking the conditions (16). In other words, we need to find t_{crit} for the given $\alpha/2$, then compare our $\bar{x}$ as per (16), and if one of the conditions is satisfied, we reject the null hypothesis and pass the t-test.

References

[1] Ilya Schurov. *Confidence intervals*. Faculty of Computer Science, Higher School of Economics. URL: https://smartedu.hse.ru/mod/page/0/798193.

[2] Ilya Schurov. *Introducing p-value*. Faculty of Computer Science, Higher School of Economics. URL: https://smartedu.hse.ru/mod/page/0/798133.

[3] Ilya Schurov. *One-sample Student's t-test*. Faculty of Computer Science, Higher School of Economics. URL: https://smartedu.hse.ru/mod/page/0/798143.

[4] Ilya Schurov. *Properties of CDF*. Faculty of Computer Science, Higher School of Economics. URL: https://smartedu.hse.ru/mod/page/0/756818.

[5] Ilya Schurov. *Type I and type II errors*. Faculty of Computer Science, Higher School of Economics. URL: https://smartedu.hse.ru/mod/page/0/798130.