12. Small-Sample Test about a Mean. Large-Sample Test about a Proportion

Small-sample test about mean: t-test about μ

The t-test about μ is employed when one wants to check a claim about the population mean, and the sample size n is small, typically, $n < 30$.

In contrast to the z-test about μ, here we have an additional restriction on the usage of the test. The t-test can be employed only if we believe that the population has normal distribution. This condition can be checked via a box plot method. If the box plot of the sample is fairly symmetric and there are no outliers, we can claim that data came from a normal population. We also can run a normality test.

t-test about μ: setup

The null hypothesis:

$$H_0: \mu = \mu_0$$

Possible alternatives:

$$H_a: \mu > \mu_0$$

$$H_a: \mu < \mu_0$$

$$H_a: \mu \neq \mu_0$$

Which type of alternative and what test value μ_0 will be used is determined by the problem context.

t-test about μ: test statistic

To test the research hypothesis we use the following test statistic

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

If H_0 is true then one can show that random variable t has t-distribution with $n - 1$ degrees of freedom. We use this fact to make our judgment whether the sample is consistent with the null or the alternative.

t-test about μ: rejection region

Three possible expressions for the rejection region:

- **If we test $H_a: \mu > \mu_0$, then $RR = [t_{\alpha,n-1}, +\infty]$**
- **If we test $H_a: \mu < \mu_0$, then $RR = [-\infty, -t_{\alpha,n-1}]$**
- **If we test $H_a: \mu \neq \mu_0$, then $RR = \left[-\infty, -t_{\alpha/2,n-1}\right] \cup [t_{\alpha/2,n-1}, +\infty]$**

The rejection region tells us which values of the test statistic t are not consistent with the null hypothesis H_0 when it is tested versus a given alternative H_a.

t-test about μ: conclusion based on rejection region

As before, the decision depends on the relationship between the rejection region and the test statistic. The same universal rule is used.

- If the test statistic falls *inside* the rejection region we *accept* H_a, the research claim.
- If the test statistic falls *outside* the rejection region we *reject* H_a.

t-test about μ: p-value remark

The p-value approach also can be used. In fact, if we have access to statistical software (for instance, MINITAB), then it is a preferred method. But since our table of critical values of t-distribution contains only seven possible choices for a critical level, in class we always will use the rejection region approach.

Example

Example 1. A federal regulatory agency is investigating an advertised claim that a certain device can increase the gasoline mileage of cars. Seven such devices are purchased and installed in seven cars belonging to the agency. Gasoline mileage change for each of the cars under standard conditions is recorded. The sample mean is .5 and the sample standard deviation of the change is 3.77. Does these data support the advertised claim? Assume that data are normally distributed. Use $\alpha = .05$.

1. Setup

Let μ be the mean gasoline mileage change for all the cars that use this device.

$$H_0: \mu = 0$$

$$H_a: \mu > 0$$

2. Test statistic

The test statistic is equal to

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{.5 - 0}{3.77/\sqrt{7}} \approx .35$$

If H_0 is true the test statistic has t-distribution with 6 d.f.

3. Rejection region

The rejection region is given by

$$RR = \left[t_{\alpha,n-1}, +\infty\right] = \left[t_{.05,6}, +\infty\right] = [1.943, +\infty]$$

4. Conclusion

The test statistic does not fall into the rejection region. Therefore, we reject the alternative at significance level of 5%. That is, our data does not support the advertised claim.

Large-sample test about population proportion: z-test about p

The z-test about p is employed when one wants to check a claim about the population proportion, and the sample size n is large, typically, $np(1-p) > 5$.

z-test about p: setup

The null hypothesis:

$$H_0: p = p_0$$

Possible alternatives:

$$H_a: p > p_0$$

$$H_a: p < p_0$$

$$H_a: p \neq p_0$$

Which type of alternative and what test value p_0 will be used is determined by the problem context.

z-test about p: test statistic

To test the research hypothesis we use the following test statistic

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

If H_0 is true then by the central limit theorem random variable z has approximately standard normal distribution. We use this fact to make our judgment whether the sample is consistent with the null or the alternative.

z-test about p: rejection region

Three possible expressions for the rejection region:

- If we test $H_a: p > p_0$, then $RR = [z_\alpha, +\infty]$
- If we test $H_a: p < p_0$, then $RR = [-\infty, -z_\alpha]$
- If we test $H_a: p \neq p_0$, then $RR = [-\infty, -z_{\alpha/2}] \cup [z_{\alpha/2}, +\infty]$

The rejection region tells us which values of the test statistic z are not consistent with the null hypothesis H_0 when it is tested versus a given alternative H_a.

z-test about p: conclusion based on rejection region

The decision depends on the relationship between the rejection region and the test statistic. The universal rule is

- If the test statistic falls *inside* the rejection region we *accept* H_a, the research claim.
- If the test statistic falls *outside* the rejection region we *reject* H_a.

z-test about p: p-value method

Instead of the rejection region approach we can calculate p-value. There are three formulas depending on a type of the alternative.

- If we test $H_a: p > p_0$, then p-value $= P(Z > z)$
- If we test $H_a: p < p_0$, then p-value $= P(Z < z)$
- If we test $H_a: p \neq p_0$, then p-value $= 2 \times P(Z > |z|)$

p-value is the probability of obtaining a test statistic value at least as extreme as our observed test statistic, assuming that H_0 is true. If this probability is small it means that our observed test statistic is too extreme to be consistent with H_0.

z-test about p: conclusion based on p-value

The universal rule is

- If p-value $< \alpha$, then we *accept* H_a.
- If p-value $> \alpha$, then we *reject* H_a.

z-test about p: example

Example 2. A market research firm interviewed 1586 potential automobile buyers by phone. One of the several questions asked was whether the customer would prefer a passenger side airbag or $300 discount on the price of the car. 847 customers prefer the airbag. Formulate and test at the significance level of 1% the null hypothesis that customers are equally divided in preferring the airbag or the discount.

1. Setup

Let p be the true proportion of all the potential customers preferring the airbag.

$$H_0 : p = .5$$

$$H_a : p \neq .5$$

2. Test statistic

Test statistic is given by

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{847/1586 - .5}{\sqrt{.5 \times .5/1586}} \approx 2.7$$

3. Rejection region

The rejection region is given by

$$
\begin{aligned}
\boldsymbol{RR} &= \left[-\infty, -\boldsymbol{z}_{\boldsymbol{\alpha/2}}\right] \cup \left[\boldsymbol{z}_{\boldsymbol{\alpha/2}}, +\infty\right] \\
&= \left[-\infty, -\boldsymbol{z}_{.\mathbf{005}}\right] \cup \left[\boldsymbol{z}_{.\mathbf{005}}, +\infty\right] \\
&= \left[-\infty, -\mathbf{2.58}\right] \cup \left[\mathbf{2.58}, +\infty\right]
\end{aligned}
$$

4. Conclusion

Since the test statistic falls inside the rejection region we accept H_a at significance level of 1%, that is, the customers are *not* equally divided in preferring the airbag or the discount.

3`. p-value

The p-value of the test is

p-value $= 2 \times P(Z > 2.7) = 2 \times .0035 = .0070$

4`. Conclusion based on p-value

Since p-value $=.007 < \alpha =.01$, we accept H_a, that is, the customers are *not* equally divided in preferring the airbag or the discount.

One-sample testing: exercises

Exercise 1. "Take the Pepsi Challenge" was a marketing campaign used by the Pepsi-Cola Company. Coca-Cola drinkers participated in a blind taste test where they were asked to taste unmarked cups of Pepsi and Coke and were asked to select their favorite. In one Pepsi television commercial, an announcer states that "in recent blind taste tests, more than half the Diet Coke drinkers surveyed said they preferred the taste of Diet Pepsi" (Consumer's Research, May 1993). Suppose 100 Diet Coke drinkers took the Pepsi Challenge and 56 preferred the taste of Diet Pepsi. Test the hypothesis that more than half of *all* Diet Coke drinkers will select Diet Pepsi in the blind taste test. Use $\alpha = .05$. What are the consequences of the test results from Coca-Cola's perspective?

One-sample testing: exercises

Exercise 2. The Lincoln Tunnel (under the Hudson River) connects suburban New Jersey to midtown Manhattan. On Mondays at 8:30 A.M., the mean number of cars waiting in line to pay the Lincoln Tunnel toll is 1220. Because of the substantial wait during rush hour, the Port Authority of New York and New Jersey is considering raising the amount of the toll between 7:30 and 8:30 A.M. to encourage more drivers to use the tunnel at an earlier or later time (Newark Star-Ledge; Aug. 27, 1995). Suppose the Port Authority experiments with peak-hour pricing for six months, increasing the toll from $4 to $7 during the rush hour peak.

One-sample testing: exercises

On 10 different workdays at 8:30 A.M. aerial photographs of the tunnel queues are taken and the number of vehicles counted. The results follow:

1260, 1052, 1201, 942, 1062, 999, 931, 849,867, 735

That is, we have: $n = 10$, $\bar{X} = 989.8$, and $s = 160.7$

Analyze the data for the purpose of determining whether peak-hour pricing succeeded in reducing the average number of vehicles attempting to use the Lincoln Tunnel during the peak rush hour.

One-sample testing: exercises

Exercise 3. In order to be effective, the mean length of life of a certain mechanical component used in a spacecraft must be larger than 1,100 hours. Owing to the prohibitive cost of this component, only three can be tested under simulated space conditions. The lifetimes (hours) of the components were recorded and the following statistics were computed: $\overline{X} = 1173.6$ and $s = 36.3$. Would you recommend that this component be passed as meeting specifications? Test using $\alpha = .05$.

One-sample testing: exercises

Exercise 4. Sales promotions that are used by manufacturers to entice retailers to carry, feature, or push the manufacturer's products are called trade promotions. A survey of 250 manufacturers conducted by Cannondale Associates, a sales and marketing consulting firm, found that 55 % of the manufacturers believe their spending for trade promotions is inefficient (Potentials in Marketing, June 1995). Is this sufficient evidence to support a previous claim by the American Marketing Association that no more than half of all manufacturers are dissatisfied with their trade promotion spending? Conduct the appropriate hypothesis test at $\alpha = .02$.

One-sample testing: exercises

Exercise 5. A consumer protection group is concerned that a ketchup manufacturer is filling its 20-ounce family-size containers with less than 20 ounces of ketchup. The group purchases 10 family-size bottles of this ketchup, weighs the contents of each, and finds that the mean weight is equal to 19.86 ounces, and the standard deviation is equal to .22 ounce. Do the data provide sufficient evidence for the consumer group to conclude that the mean fill per family-size bottle is less than 20 ounces? Test using $\alpha =.05$.

One-sample testing: exercises

Exercise 6. The manufacturer of an over-the-counter analgesic claims that its product brings pain relief to headache sufferers in less than 3.5 minutes, on average. In order to be able to make this claim in its television advertisements, the manufacturer was required by a particular television network to present statistical evidence in support of the claim. The manufacturer reported that for a random sample of 50 headache sufferers, the mean time to relief was 3.3 minutes and the standard deviation was 1.1 minutes. Do these data support the manufacturer's claim? Test using $\alpha = .05$.