Anna Janicka

Mathematical Statistics 2020/2021
Lecture 10

1. Hypothesis Testing – Examples of LR tests

In the previous lecture, we looked at the general framework of designing likelihood ratio tests. Now we will look at some examples of LR tests most commonly used in practice.

1.1. **Single population.** We will start by constructing tests which are used to verify claims that a value of a distribution parameter is equal to something (against the alternative that it is equal to something else), i.e. in cases where we have observations for a single population (one sample) and we want to say something about the value of the parameters for this single population (or sample).

1.1.1. *Model I.* Lest us first assume that we have a random sample $X_1, \ldots, X_n$ from a normal distribution with parameters μ and σ^2, where σ^2 **is known**. Let us assume that we want to test the null hypothesis that $\mu = \mu_0$, against different types of alternatives. Noting the examples discussed in the previous lecture (when we constructed the LR test for the mean in a normal model), we see that we can use the test statistic

$$U = \frac{\bar{X} - \mu_0}{\sigma}\sqrt{n},$$

which under the null hypothesis has a distribution $N(0,1)$, to construct critical regions in the following way:

- If the alternative is that $\mu > \mu_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : U(x) > u_{1-\alpha}\}$$
- If the alternative is that $\mu < \mu_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : U(x) < -u_{1-\alpha} = u_\alpha\}$$
- If the alternative is that $\mu \neq \mu_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : |U(x)| > u_{1-\alpha/2}\},$$

where u_p signifies the quantile of rank p of the standard normal distribution.

Example:

(1) Suppose we have a random sample of 10 observations from a normal distribution with an unknown mean μ and variance equal to 1:

-1.21 -1.37 0.51 0.37 -0.75 0.44 1.20 -0.96 -1.14 -1.40

We wish to verify the null hypothesis that the mean is equal to 0, against the alternative that it is not, at a significance level $\alpha = 0.05$.

The sample average amounts to -0.43, so the value of the appropriate test statistic amounts to $U = \frac{-0.43-0}{1}\sqrt{10} \approx -1.36$. The critical value for a two-sided test at the $\alpha = 0.05$ significance level amounts to $u_{0.975} \approx 1.96$. The value of the test statistic does not fall into the critical region of $(-\infty, -1.96) \cup (1.96, \infty)$, so we do not have grounds to reject the null. We could also note that the p-value of the result amounts to $2 \cdot \Phi(-1.36) \approx 2 \cdot 0.086 = 0.172$. Since the p-value is larger than the adopted significance level, we do not have grounds to reject the null.

If we wished to verify the null hypothesis that the mean is equal to 0, against the alternative that it is smaller, at the same significance level of $\alpha = 0.05$, the result would stay the same: the value of the test statistic -1.36 does not fall into the critical region for the one-sided test, which is equal to $(-\infty, -1.64 = u_{0.05})$. However, since the p-value of the result is now equal to $\Phi(-1.36) \approx 0.086$, we would have rejected the null in favor of the alternative at a significance level of $\alpha = 0.1$.

1.1.2. *Model II.* Let us now assume that we still have a random sample $X_1, \ldots, X_n$ from a normal distribution with parameters μ and σ^2, but now σ^2 **is unknown**. Let us assume that just like before, we want to test the null hypothesis that $\mu = \mu_0$, against different types of alternatives. Similarly to the previous case, we can use a test statistic

$$T = \frac{\bar{X} - \mu_0}{S}\sqrt{n},$$

which under the null hypothesis has a distribution $t(n-1)$, to construct critical regions in the following way:

- If the alternative is that $\mu > \mu_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : T(x) > t_{1-\alpha}(n-1)\}$$
- If the alternative is that $\mu < \mu_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : T(x) < -t_{1-\alpha} = t_\alpha(n-1)\}$$
- If the alternative is that $\mu \neq \mu_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : |T(x)| > t_{1-\alpha/2}(n-1)\},$$

where $t_p(n-1)$ is the quantile of rank p of the t-Student distribution with, $n-1$ degrees of freedom, and S^2 is the unbiased estimator of the variance.

Since in this case we do not know the value of σ^2, we might also want to verify hypotheses for this parameter. Let us assume that we want to test $\sigma = \sigma_0$, against different types of alternatives. In this case, we can use a test statistic

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2},$$

which under the null hypothesis has a $\chi^2(n-1)$ distribution, to construct critical regions in the following way:

- If the alternative is that $\sigma > \sigma_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : \chi^2(x) > \chi^2_{1-\alpha}(n-1)\}$$
- If the alternative is that $\sigma < \sigma_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : \chi^2(x) < \chi^2_\alpha(n-1)\}$$
- If the alternative is that $\sigma \neq \sigma_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : \chi^2(x) > \chi^2_{1-\alpha/2}(n-1) \vee \chi^2(x) < \chi^2_{\alpha/2}(n-1)\},$$

where $\chi^2_p(n-1)$ is the quantile of rank p of the chi-square distribution with $n-1$ degrees of freedom, and S^2 is the unbiased estimator of the variance.

Example:

(1) Suppose again we have a random sample of 10 observations from a normal distribution, but this time we do not know neither the mean μ nor the variance:

-1.21 -1.37 0.51 0.37 -0.75 0.44 1.20 -0.96 -1.14 -1.40

We wish to verify the null hypothesis that the mean is equal to 0, against the alternative that it is not, at a significance level $\alpha = 0.05$.

The sample average amounts to -0.43, and the sample variance amounts to 0.92, so the value of the appropriate test statistic amounts to $T = \frac{-0.43-0}{\sqrt{0.92}}\sqrt{10} \approx -1.42$. The critical value for a two-sided test at the $\alpha = 0.05$ significance level amounts to $t_{0.975}(9) \approx 2.26$. Since the value of the test statistic does not fall into the critical region of $(-\infty, -2.26) \cup (2.26, \infty)$, we do not have grounds to reject the null. We could also note that the p-value of the result amounts to 0.188. Since the p-value is larger than the adopted significance level, we do not have grounds to reject the null.

If we wished to verify the null hypothesis that the mean is equal to 0, against the alternative that it is smaller, at the same significance level of $\alpha = 0.05$, the result would stay the same: the value of the test statistic -1.42 does not fall into the critical region for the one-sided test, which is equal to $(-\infty, -1.83 = t_{0.05}(9))$. However, since the p-value of the result is now equal to ≈ 0.094, we would have rejected the null in favor of the alternative at a significance level of $\alpha = 0.1$.

Now, if we also wished to verify whether the variance of the distribution is in fact equal to 1, against the alternative that it is not, at a significance level $\alpha = 0.05$, we would have used a test statistic equal to $\chi^2 = \frac{9 \cdot 0.92}{1} \approx 8.28$. Since the value of the test statistic does not fall into the critical region $(0, \chi^2_{0.025}(9) \approx 2.70) \cup (\chi^2_{0.975}(9) \approx 19.02, \infty)$, we do not have grounds to reject the null hypothesis in favor of the two-sided alternative.

However, if our wish was to verify the null that $\sigma^2 = 7$ against the alternative that it is smaller, we would have used a test statistic equal to $\chi^2 = \frac{9 \cdot 0.92}{\sqrt{7}} \approx 3.12$. Since the value of the test statistic falls into the critical region $(0, \chi^2_{0.05}(9) \approx 3.33)$, we should reject the null that $\sigma^2 = 7$ in favor of the alternative that in fact the variance is smaller.

1.1.3. *Model III.* Lest us now consider the case where e have a random sample $X_1, \ldots, X_n$ from a distribution which is not normal, but we have a large sample size (n is large enough for the CLT to provide a good approximation). Let us assume that we want to test the null hypothesis that the mean of this distribution $\mu = \mu_0$, against different types of alternatives. In this case, we might use a test with a critical region which has approximately the required significance level α, by taking

$$U = \frac{\bar{X} - \mu_0}{S}\sqrt{n},$$

which under the null hypothesis has (approximately, for large n) a standard normal distribution, to construct critical regions in the following way:

- If the alternative is that $\mu > \mu_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : U(x) > u_{1-\alpha}\}$$
- If the alternative is that $\mu < \mu_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : U(x) < -u_{1-\alpha} = u_\alpha\}$$
- If the alternative is that $\mu \neq \mu_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : |U(x)| > u_{1-\alpha/2}\},$$

where u_p signifies the quantile of rank p of the standard normal distribution and S^2 is the unbiased estimator of the variance.

1.1.4. *Model IV.* As a special case of model III, we might consider the situation when the random variables $X_1, \ldots, X_n$ that we observe come from a distribution such that $P(X = 1) = p = 1 - P(X = 0)$. In this case, the mean of the distribution is equal to the probability p, and we might want to make use of the de Moivre-Laplace theorem to construct an asymptotic test. If we have a large sample size (n is large enough for the CLT to provide a good approximation) and we want to test the null hypothesis that $p = p_0$, against different types of alternatives, we might use the test statistic provided in Model III. It is not optimal, however, since it requires calculating the variance for the sample of observations. We might prefer to use a more comfortable estimator for the variance instead, and take:

$$U^* = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)}}\sqrt{n},$$

which under the null hypothesis has (approximately, for large n) a standard normal distribution, and construct critical regions in the following way:

- If the alternative is that $p > p_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : U^*(x) > u_{1-\alpha}\}$$
- If the alternative is that $p < p_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : U^*(x) < -u_{1-\alpha} = u_\alpha\}$$
- If the alternative is that $p \neq p_0$, then the critical region of the test for significance level α is equal to
$$C^* = \{x : |U^*(x)| > u_{1-\alpha/2}\},$$

where u_p signifies the quantile of rank p of the standard normal distribution.

Example:

(1) We toss a coin 400 times and we observe 180 heads. Is the coin symmetric?

We will be testing the null that $p = \frac{1}{2}$, so we will use a test statistic of the form $U^* = \frac{\frac{180}{400} - \frac{1}{2}}{\sqrt{\frac{1}{2}(1-\frac{1}{2})}}\sqrt{400} = -2$. If we wish to test the null against the two-sided alternative, at a significance level $\alpha = 0.05$, we should use a critical region of the form $(-\infty, -1.96 = -u_{0.975}) \cup (1.96 = u_{0.975}, \infty)$. Since the value of the test statistic falls into this critical region, we should reject the null – we can't reasonably assume that the coin is symmetric, at a significance level 0.05.