Listy Biometryczne - Biometrical Letters
Vol. 40 (2003), No. 1, 9-14

A note on testing mean in normal population

Joanna Tarasińska

Department of Applied Mathematics, Agricultural University,
Akademicka 13, 20-934 Lublin, Poland
johata@ursus.ar.lublin.pl

SUMMARY

The paper concerns the problem of testing the expected value of normal distribution by means of Student's t-test. Choosing the form of alternative hypothesis (both-, left- or right-sided) one shouldn't take into consideration the measurements in the sample. In such a case applied test has got less power than assumed.

KEY WORDS: normal distribution, Student's t-test, power of the test.

1. Introduction

Let $X \sim N(\mu, \sigma^2)$, with σ unknown, and let $X_1, X_2, ..., X_n$ be the random sample. Let us consider the null hypothesis $H_0 : \mu = \mu_0$ versus H_1 ($\mu \neq \mu_0$ or $\mu > \mu_0$ or $\mu < \mu_0$). The t-test based on test statistic

$$t = \frac{\overline{X} - \mu_0}{\sqrt{\dfrac{nS^2}{n(n-1)}}} \tag{1}$$

is used, where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $nS^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2$. It is known that when we choose the form of the alternative H_1 we shouldn't take into consideration the measurements in the sample. On the contrary, the hypothesis H_1 should be fixed before taking the sample. However, the inclination to being influenced by the value of arithmetic mean $\overline{X}$ when H_1 is to be stated is often observable, especially in the case of practitioners, non statisticians and students being taught statistics.

The aim of the paper is to demonstrate the influence of such an incorrect approach on the quality of the test.

The case when σ is known is also considered.

2. The significance level and the power of test, σ unknown

Let us assume that somebody establishes the form of the alternative H_1 as follows:

– if $\overline{X} < \mu_0$ then H_1 is left-sided ($H_1 : \mu < \mu_0$),

– if $\overline{X} > \mu_0$ then H_1 is right-sided ($H_1 : \mu > \mu_0$),

and tests H_0 versus H_1 rejecting H_0 when the value of the test statistic is in the interval:

– $(-\infty, t_\alpha)$ in case $H_1 : \mu < \mu_0$,

– $(t_{1-\alpha}, \infty)$ in case $H_1 : \mu > \mu_0$,

where t_p is the p-th quantile of the Student's t distribution with $n-1$ degrees of freedom (t_{n-1}).

If anyone who applied the t-test in such a way as described above thinks he tests H_0 versus H_1 on significance level α, he is in a mistake. It is not difficult to prove that significance level is in fact 2α. Indeed, the significance level can be written as $P(t < t_\alpha, U < 0) + P(t > t_{1-\alpha}, U > 0)$, where $t = \frac{U\sqrt{n-1}}{\sqrt{V}} \sim t_{n-1}$ and $U = \frac{(\overline{X}-\mu_0)\sqrt{n}}{\sigma}$, $V = \frac{nS^2}{\sigma^2}$ are independent random variables distributed as, respectively, $N(0,1)$ (i.e., standard normal) and χ^2_{n-1} (chi-square with $n-1$ degrees of freedom). So, the significance level is

$$P(t < t_\alpha, U < 0) + P(t > t_{1-\alpha}, U > 0) = P(t < t_\alpha) + P(t > t_{1-\alpha}) = 2\alpha.$$

If we want to have significance level α quaranteed, we have to modify the considered test as follows: reject H_0 when the value of test statistic is in:

– $\left(-\infty, t_{\frac{\alpha}{2}}\right)$ in case $H_1 : \mu < \mu_0$;

– $\left(t_{1-\frac{\alpha}{2}}, \infty\right)$ in case $H_1 : \mu > \mu_0$.

Let us call the considered test on significance level α as the "conditional" test as it is conditioned by the value of $\overline{X}$. Let us see how the power function of it looks like.

If $\mu = \mu_0 + k\sigma$, $k > 0$, then the test function (1) has noncentral t distribution with $n-1$ degrees of freedom and the noncentrality parameter $\lambda = k\sqrt{n}$ (Patel et al., 1976). We can write it in the form

$$t = (U + k\sqrt{n})\sqrt{n-1} \cdot \frac{1}{\sqrt{V}},$$

where $U = \frac{(\overline{X}-\mu_0)\sqrt{n}}{\sigma} - k\sqrt{n} \sim N(0,1)$. So, the power of the test (the probability of rejecting H_0 when H_1 is true) is

$$M(k) = P\left(t > t_{1-\frac{\alpha}{2}}, \overline{X} > \mu_0\right) = 1 - G_{n-1,k\sqrt{n}}\left(t_{1-\frac{\alpha}{2}}\right),$$

where $G_{v,\lambda}$ is the cumulative distribution function of the noncentral t distribution with v degrees of freedom and the parameter of noncentrality λ.

For $\mu = \mu_0 + k\sigma, k < 0$, we have in the same way

$$M(k) = G_{n-1,k\sqrt{n}}\left(t_{\frac{\alpha}{2}}\right).$$

Let us notice now that the considered test is not unbiased (Lehmann, 1986) because $M(k) \xrightarrow{k\to 0} \frac{\alpha}{2}$.

We want to compare the power of the considered "conditional" test and the power of "standard" test $H_0 : \mu = \mu_0$ *versus* $H_1 : \mu \neq \mu_0$, both on the significance level α. The power of the "standard" test is equal to

$$M_s(k) = P\left(t > t_{1-\frac{\alpha}{2}}\right) + P\left(t < t_{\frac{\alpha}{2}}\right) = 1 - G_{n-1,k\sqrt{n}}\left(t_{1-\frac{\alpha}{2}}\right) + G_{n-1,k\sqrt{n}}\left(t_{\frac{\alpha}{2}}\right).$$

Figure 1 presents both powers for $\alpha = 0.05$, $n = 10$ and $-0.5 \leq k \leq 0.5$. The cdf of noncentral t distribution was calculated using the formula given by Owen (1968, p. 464). Computations were made by means of a programme written in Maple V Release. For bigger $|k|$ the curves coincide. Table 1 presents quotients $\frac{M(k)}{M_s(k)}$ for $\alpha = 0.05$ and $\alpha = 0.01$, $k = 0(0.1)0.5$, $n = 3(1)20$.

Table 1. The quotients $\dfrac{M(k)}{M_s(k)}$, σ unknown

	$\alpha = 0.05$					$\alpha = 0.01$				
k	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
n										
3	0.632	0.746	0.835	0.898	0.938	0.634	0.751	0.840	0.902	0.942
4	0.668	0.802	0.891	0.943	0.971	0.676	0.813	0.901	0.950	0.976
5	0.697	0.841	0.925	0.966	0.985	0.711	0.858	0.937	0.974	0.989
6	0.722	0.871	0.946	0.979	0.992	0.741	0.891	0.959	0.986	0.995
7	0.743	0.893	0.960	0.986	0.995	0.766	0.915	0.973	0.992	0.997
8	0.761	0.910	0.970	0.991	0.997	0.788	0.932	0.981	0.995	0.999
9	0.777	0.924	0.977	0.993	0.998	0.806	0.946	0.986	0.997	0.999
10	0.791	0.935	0.982	0.995	0.999	0.823	0.956	0.990	0.998	1.000
11	0.804	0.944	0.986	0.997	0.999	0.837	0.964	0.993	0.999	1.000
12	0.816	0.952	0.989	0.998	0.999	0.850	0.970	0.995	0.999	1.000
13	0.826	0.958	0.991	0.988	1.000	0.861	0.975	0.996	0.999	1.000
14	0.836	0.963	0.993	0.999	1.000	0.871	0.979	0.997	1.000	1.000
15	0.845	0.967	0.994	0.999	1.000	0.880	0.982	0.998	1.000	1.000
16	0.853	0.971	0.995	0.999	1.000	0.889	0.985	0.998	1.000	1.000
17	0.860	0.974	0.996	0.999	1.000	0.896	0.987	0.998	1.000	1.000
18	0.867	0.977	0.997	0.999	1.000	0.903	0.989	0.999	1.000	1.000
19	0.874	0.980	0.997	1.000	1.000	0.909	0.990	0.999	1.000	1.000
20	0.880	0.982	0.998	1.000	1.000	0.915	0.991	0.999	1.000	1.000

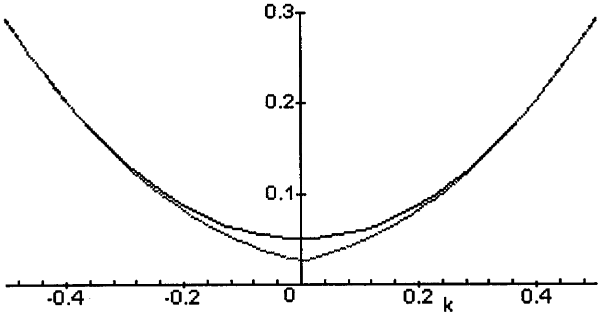

Figure 1. The power of the "standard" test (the upper curve) and the "conditional" one (the lower curve), $\alpha = 0.05$, $n = 10$, σ unknown

3. Case σ known

If the standard deviation of the random variable X is known then the "conditional" test on significance level α rejects $H_0 : \mu = \mu_0$ when the value of the test function $U = \frac{(\overline{X}-\mu_0)\sqrt{n}}{\sigma}$ is in:

- $\left(-\infty,\ u_{\frac{\alpha}{2}}\right)$, in case $\overline{X} < \mu_0 \quad (H_1 : \mu < \mu_0)$,
- $\left(\ \mu_{1-\frac{\alpha}{2}},\ \infty\right)$, in case $\overline{X} > \mu_0 \quad (H_1 : \mu > \mu_0)$,

where u_p is the pth quantile of the standard normal distribution. The power of the "conditional" test is

$$M(k) = \begin{cases} 1 - \Phi(\mu_{1-\frac{\alpha}{2}} - k\sqrt{n}) & \text{for } k > 0, \\ \Phi(u_{\frac{\alpha}{2}} - k\sqrt{n}) & \text{for } k < 0, \end{cases}$$

where $\mu = \mu_0 + k\sigma$ and $\Phi(\cdot)$ is the cdf of standard normal distribution. The power of the "standard" test is $M_s(k) = 1 - \Phi(u_{1-\frac{\alpha}{2}} - k\sqrt{n}) + \Phi(u_{\frac{\alpha}{2}} - k\sqrt{n})$. Table 2 presents quotients $\dfrac{M(k)}{M_s(k)}$ for $\alpha = 0.05$ and $\alpha = 0.01$, $k = 0(0.1)0.5$, $n = 3(1)20$.

4. Conclusions

The procedure of testing should be applied in a proper way, which means that the hypotheses should be fixed before taking the sample. If anybody conditions the form of the alternative H_1 by the result of measurements he ought to know that his test has got less power.

Table 2. The quotients $\frac{M(k)}{M_s(k)}$, σ known

	$\alpha = 0.05$					$\alpha = 0.01$				
k	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
n										
3	0.692	0.835	0.919	0.963	0.983	0.731	0.881	0.953	0.982	0.993
4	0.718	0.867	0.943	0.977	0.991	0.761	0.910	0.970	0.990	0.997
5	0.740	0.890	0.959	0.985	0.995	0.785	0.930	0.980	0.994	0.998
6	0.759	0.908	0.969	0.990	0.997	0.805	0.945	0.986	0.997	0.999
7	0.775	0.923	0.976	0.993	0.998	0.822	0.955	0.990	0.998	1.000
8	0.790	0.934	0.982	0.995	0.999	0.837	0.964	0.993	0.999	1.000
9	0.803	0.943	0.986	0.996	0.999	0.850	0.970	0.995	0.999	1.000
10	0.814	0.951	0.988	0.997	0.999	0.862	0.975	0.996	0.999	1.000
11	0.825	0.957	0.991	0.998	1.000	0.872	0.979	0.997	1.000	1.000
12	0.835	0.963	0.992	0.999	1.000	0.881	0.982	0.998	1.000	1.000
13	0.844	0.967	0.994	0.999	1.000	0.890	0.985	0.998	1.000	1.000
14	0.852	0.971	0.995	0.999	1.000	0.897	0.987	0.999	1.000	1.000
15	0.860	0.974	0.996	0.999	1.000	0.904	0.989	0.999	1.000	1.000
16	0.867	0.977	0.996	0.999	1.000	0.910	0.990	0.999	1.000	1.000
17	0.873	0.980	0.997	1.000	1.000	0.916	0.992	0.999	1.000	1.000
18	0.879	0.982	0.998	1.000	1.000	0.921	0.993	0.999	1.000	1.000
19	0.885	0.984	0.998	1.000	1.000	0.926	0.994	0.999	1.000	1.000
20	0.890	0.985	0.998	1.000	1.000	0.930	0.994	1.000	1.000	1.000

It can be seen from Tables 1 and 2 that it is especially acute for small deviations of μ from μ_0 and small size of the sample. The "conditional" test can have even dozen percent less power than the "standard" one on the same significance level, though the powers of both tests in such a case are rather small. The differences of powers are bigger for $\alpha = 0.05$ than for $\alpha = 0.01$ and decrease when the sample size increases.

Finally, note that the differences between powers of the "standard" and "conditional" tests are smaller when σ is known than when σ is unknown.

REFERENCES

Lehmann E.L.(1986). *Testing statistical hypotheses.* New York, Wiley.

Owen D.B.(1968). A survey of properties and applications of the noncentral t-distribution. *Technometrics* **10**, 445-478.

Patel J.K., Kapadia C.H., Owen D.B. (1976). *Handbook of statistical distributions. Marcel Dekker Inc.*, New York.

Received 25 February 2003

Uwagi o testowaniu średniej w populacji normalnej

STRESZCZENIE

W pracy rozważa się problem testowania hipotez dotyczących wartości oczekiwanej rozkładu normalnego testem t-Studenta. Wybierając hipotezę alternatywną (lewo-, prawo lub obustronną) nie należy brać pod uwagę wyników otrzymanych w próbie. W przeciwnym razie zastosowany test ma moc mniejszą, niż przypuszczamy.

SŁOWA KLUCZOWE: rozkład normalny, test t-Studenta, moc testu.