JEL 378.14: 519.8

TAIL PROBABILITIES CONFIDENCE BOUNDS

Kateryna Vadymivna Stiepanova, Ph.D.,
Simon Kuznets Kharkiv National University of Economics, Kharkiv, Ukraine

Abstract — **The aim of this brief communication is to compare the calculations of confidence bounds based on the same data but using different methods.**

Key Terms — **Confidence bounds, noncentral t-distribution, tail probability.**

The noncentral t-distribution is intimately tied to statistical inference procedures for samples from normal populations. For simple random samples from a normal population the usage of the noncentral t-distribution includes basic power calculations, variables acceptance sampling plans and confidence bounds for quantiles, tail probabilities, statistical process control parameters and variation coefficients. The purpose of these notes is to describe these applications in some detail, giving sufficient theoretical derivation so that these procedures may easily be extended to more complex normal data structures, that occur, for example, in multiple regression and analysis of variance settings. Firs of all, we are going to give a definition of the noncentral t-distribution, that ties directly into all the applications. This is demonstrated upfront by exhibiting the basic probabilistic relationship underlying all these applications. Note, that a short list [1]-[6] is provided to give an entry into the relevant literature.

So, let Z and V are independent standard normal and chi-square random variables respectively, the latter with f degrees of freedom, then the following ratio have a noncentral t-distribution

$$T_{f,\delta} = \frac{Z+\delta}{\sqrt{V/f}}$$

with f degrees of freedom and noncentrality parameter δ. Although $f \geq 1$ originally was intended to be an integer closely linked to sample size, it is occasionally useful to extend its definition to any real $f > 0$. The noncentrality parameter δ may also be any real number. The cumulative distribution function (briefly cdf) of $T_{f,\delta}$ is denoted by

$$P_{f,\delta}(t) = P(T_{f,\delta} \leq t).$$

Remark here, that if $\delta = 0$, then the noncentral t-distribution reduces to the usual central or Student t-distribution $G_{f,\delta}(t)$ increases from 0 to 1 as t increases from $-\infty$ to $+\infty$ and it decreases from 1 to 0 as δ increases from $-\infty$ to $+\infty$. While the former is a standard property of any cdf, the latter becomes equally obvious when rewriting $G_{f,\delta}(t)$ as follows:

$$G_{f,\delta}(t) = P(T_{f,\delta} \leq t) = P\left(\frac{Z+\delta}{\sqrt{V/f}} \leq t\right) = $$
$$= P(Z - t\sqrt{V/f} \leq -\delta).$$

As f gets very large the distribution of $T_{f,\delta}$ approximates the normal distribution of $Z+\delta$ (see, for example, a detailed analysis of behavior and corresponding plots with comments in [1]).

The sample mean $\bar{X}$ and sample standard deviation S are respectively defined as:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}.$$

The following distributional facts are well-known:

a) $\bar{X}$ and S are statistically independent;

b) $\bar{X}$ is distributed like a normal random variable with mean μ and standard deviation $\sigma/\sqrt{n}$ or equivalently, $Z = \sqrt{n}(\bar{X}-\mu)/\sigma$ has a standard normal distribution (it means that μ=0 and standard deviation S=1);

c) $V = (n-1)S^2/\sigma^2$ has a chi-square distribution with $f = n - 1$ degrees of freedom and is statistically independent of Z.

All one-sample applications involving the noncentral t-distribution can be reduced to calculating the following probability

$$\gamma = P(\bar{X} - aS \le b). \qquad (1)$$

To relate this probability to the noncentral t-distribution note the equivalence of the following three inequalities, which can be established by simple algebraic calculations:

$$\bar{X} - aS \le b \Leftrightarrow T_{f,\delta} = \frac{Z+\delta}{\sqrt{V/f}} \le a\sqrt{n} \Leftrightarrow$$
$$\Leftrightarrow Z + \delta \le a\sqrt{nV/f}.$$

Using definition of Z, V (which was defined previously in terms of $\bar{X}$ and S), we have:

$$\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} + \delta \le a\sqrt{n(n-1)\mathrm{S}^2/\sigma^2 f}.$$

Finally, substituting $\delta = -\sqrt{n}(b-\mu)/\sigma$ and $f = n-1$, we obtain the inequality

$$\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} - \frac{\sqrt{n}(b-\mu)}{\sigma} \le$$
$$\le a\sqrt{n(n-1)\mathrm{S}^2/\sigma^2(n-1)},$$

that leads after simple algebraic calculations and reducing to

$$\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} - \frac{\sqrt{n}(b-\mu)}{\sigma} \le a\sqrt{n}\frac{S}{\sigma}$$

or

$$\frac{\sqrt{n}(\bar{X}-\mu)/\sigma - \sqrt{n}(b-\mu)/\sigma}{S/\sigma} \le a\sqrt{n}.$$

Thus, combine (1) with the latest inequality we conclude:

$$\gamma = P(T_{f,\delta} \le a\sqrt{n}) = \mathrm{G}_{f,\delta}(a\sqrt{n}).$$

Usually tree of four parameters (δ, a, n, γ) were given for computation of $\mathrm{G}_{f,\delta}(a\sqrt{n})$.

As we know, quality control deal with variables acceptance sampling plans (VASP). In a VASP the quality of items in a given sample is measured on a quantitative scale. An item is judged defective when its measured quality exceeds a certain threshold. The samples are drawn randomly from population of items. The objective is to make inferences about the proportions of defectives in the population. This leads either to an acceptance or a rejection of population quality as a whole. Speaking about a VASP it is usually assumes that measurements $X_1,...,X_n$ for a random sample of n items from a population are available and that defectiveness for any given sample item i is equivalent to $X_i < L$, where L is some given lower specification limit (see, for example, [2]). Assume that we deal with a random sample from a normal population with mean μ and standard deviation σ. The following note will be it terms of the tail probabilities of a normal population. For a given threshold value x_0 we interested in the tail probability

$$p = p(\mu,\sigma,x_0) = P_{\mu,\sigma}(X < x_0) = \Phi\left(\frac{x_0-\mu}{\sigma}\right),$$

here $\Phi(x)$ denotes the normal distribution function, and p can be interpreted in the context of VASP as the proportion of defective items in the population. Upper bounds for such probabilities p could give a producer the needed assurance of having a proportion of defectives $\le p_0$, the value used in setting up the VASP. Although $p = \Phi\left((x_0 - \bar{X})/S\right)$ is a natural but biased estimate of p the construction of confidence bounds is not so obvious (see [1]-[4].). Firstly we will discuss the relationships between upper and lower bounds for left and right tail probabilities p and $q = 1-p$. If $p_U(\gamma)$ denotes an upper bound for p with confidence level γ (i.e. $P_{\mu,\sigma}(p_U(\gamma) \ge p) = \gamma$), then we also have $P_{\mu,\sigma}(p_U(\gamma) \le p) = 1-\gamma$ for all (μ,σ); so that $p_U(\gamma)$ can also serve as a lower bound $p_L(1-\gamma) = p_U(\gamma)$ for p with confidence level $1-\gamma$. If the upper tail probability $q = 1-p$ of the normal distribution is of interest, then $1-p_U(\gamma)$ will serve as lower bound for q with confidence level and thus as an upper bound for q with confidence level $1-\gamma$. Thus it suffices to limit any further discussion to upper confidence

bounds for p. In order to find these upper bounds we will use the following result of [5]:
Lemma. *If X is a random variable with continuous and strictly increasing distribution function $F(t) = P(X \le t)$ then the random variable $U = F(X)$ is uniformly distributed over [0; 1], i.e., $P(\mathrm{U} \le u) = u$ for $0 \le u \le 1$.*
For constructing upper bounds we consider: $\sqrt{n}(x_0 - \bar{X}) / \mathrm{S} =$

$$\frac{\sqrt{n}(x_0 - \mu)/\sigma + \sqrt{n}(\mu - \bar{X})/\sigma}{S/\sigma} = T_{n-1,\delta}.$$

Here $\delta = \sqrt{n}(x_0 - \mu)/\sigma = \sqrt{n}\Phi^{-1}(p)$ is an increasing function of p; note that Z and Z^* ($Z = \sqrt{n}(\bar{X} - \mu)/\sigma = -Z^* = -\sqrt{n}(\mu - \bar{X})/\sigma$) have the same standard normal distribution. By the above Lemma the random variable

$$U = G_{n-1,\delta}\left(\sqrt{n}(x_0 - \bar{X})/\mathrm{S}\right) = G_{n-1,\delta}(T_{n-1,\delta})$$

has a uniform distribution over [0; 1]. Such a function U of the sample data and the unknown parameters is called a pivot when its distribution is completely known, as in case here. The concept of pivots is often employed in constructing confidence sets. So, $\gamma = P(U \ge 1-\gamma)$, since $G_{n-1,\delta}$ is decreasing in δ and we have

$$U = G_{n-1,\delta}\left(\sqrt{n}(x_0 - \bar{X})/\mathrm{S}\right) \ge 1-\gamma \Leftrightarrow \delta \le \delta,$$

here δ solves $G_{n-1,\delta}\left(\sqrt{n}(x_0 - \bar{X})/\mathrm{S}\right) = 1-\gamma.$

Hence, δ is an upper confidence bound for $\delta = \sqrt{n}\Phi^{-1}(p)$ with confidence level γ,

$$\delta \ge \delta \Leftrightarrow p_U(\gamma) = \Phi(\delta/\sqrt{n}) \ge \Phi(\delta/\sqrt{n}) = p,$$

p_U is the desired upper confidence bound for p with confidence level γ.

We point out that the coverage probability statement in $\gamma = P(U \ge 1-\gamma)$ holds for (μ, σ), which enter through U in two-fold form, namely through δ in $G_{n-1,\delta}$ and through the joint distribution of $\bar{X}$ and S in $\sqrt{n}(x_0 - \bar{X})/\mathrm{S}$. This means that the coverage probability is constant in μ and σ, thus equals the confidence coefficient or the minimum coverage probability $\bar{\gamma}$. The same comment applies to tolerance bounds.

To sum up, let us compare two previously discussed methods here. Confidence bounds based on the same data but using different methods are typically different. Furthermore, even if method which based on $\bar{X}$ and S is generally superior to binomial method, it can happen (as in this instance) that the bound produced by binomial method is better than the bound produced by the first one method. According to [1], for example, both upper bounds are above the true target 0.02275 but the binomial bound happens to be closer. Finally, we point out from [6] that the 95% confidence curve has to be interpreted point-wise, i.e., the probability for several such upper or lower bounds simultaneously covering their respective targets is less than 0.95.

References

1. Scholz F. Industrial Statistics / F. Scholz. – Washington: Washington EDU, 2014. – 58 p.
2. Hogg Robert V. Probability and Statistical Inference / R. V. Hogg, E. A. Tanis, D. L. Zimmerman. – Pearson: Pearson Higher Education, 2014. – 548 p.
3. Chou Y. M., Owen D. B., Borrego S. A. Lower Confidence Limits on Process Capability Indices / Y. M. Chou, D. B. Owen, S. A. Borrego // Journal of Quality Technology. – 1990. – 22. – P.223-229.
4. Kotz S., Lovelace C. R. Process Capability Indices in Theory and Practice / S. Kotz, C. R. Lovelace. – Arnold: London, 1998. – 279 p.
5. Lehmann E.L., Romano J.P. Testing Statistical Hypotheses / E.L. Lehmann, J.P. Romano. – New York: Third Edition, John Wiley & Sons, 2005. – 786 p.
6. Scholz F., Vangel M. Tolerance Bounds and Cpk Condence Bounds Under Batch Effects / F. Scholz, M. Vangel. – Birkhauser, Boston: Advances in Stochastic Models for Reliability, Quality and Safety, eds., Part IV, 1998. – P.361-379.

Author

Stiepanova Kateryna Vadymivna,
Assistant of the Department of Higher Mathematics, Economic and Mathematical Methods, KhNUE, Kharkov, Ukraine (stepanova.katerina@hneu.net).

Manuscript received 30 January 2017.

Published as submitted by the author.