A New Evidence of Two-sided Test for Normal Distribution with Restricted Parameter Space

Min Wang

Clemson University

March 14, 2012

Outline

Problem setup

- Let $X_1, \cdots, X_n$ be a random sample from an $N(\theta, \sigma^2)$ population.
- The location parameter $\theta \in \Theta$ is the parameter of interest.
- In this talk, we consider the two-sided hypothesis

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_a : \theta \neq \theta_0. \tag{1}$$

- Cases of σ^2 both known and unknown will be discussed.

When σ^2 is known,

The usual p-value of UMPU test of (1) with respect to an observation $\bar{x}$ is

$$p(\bar{x}) = P\left(|Z| > \frac{|\bar{x} - \theta_0|}{\sigma/\sqrt{n}}\right), \tag{2}$$

where Z represents a random variable of standard normal distribution. We reject the null hypothesis if the p-value $p(\bar{x}) < \alpha$ (usually $\alpha \in \{0.01, 0.05, 0.1\}$).

When σ^2 is unknown,

The usual p-value of UMPU test of (1) with respect to an observation $\bar{x}$ is

$$\tilde{p}(\bar{x}) = P\left(|T| > \frac{|\bar{x} - \theta_0|}{s_n/\sqrt{n}}\right),$$

where T represents a t-distribution with degrees of freedom $n-1$ and s_n is the sample standard deviation. We reject the null hypothesis if the p-value $\tilde{p}(\bar{x}) < \alpha$ (usually $\alpha \in \{0.01, 0.05, 0.1\}$).

Remarks

- The usual p-value of UMPU test performs well in practical situations
- Now consider the parameter space Θ is restricted and the bounds of the parameter space are known. For example, $\theta \in (a,\ b)$ with both a and b known.
  - The average weight of newborn infants.
  - The average age of retirement
  - ....
- The usual p-value does not sufficiently utilize such bounded information to make a more accurate decision.

Remarks

- It is necessary to present a measure of evidence against the null hypothesis dependent on the parameter space Θ.
- From the frequentist viewpoint, H. WANG (2005) recently proposed a modified p-value of two-sided test for the normal distribution with restricted parameter space, given by

$$r(\bar{x}) = \frac{P\left(|Z| > \frac{|\bar{x}-\theta_0|}{\sigma/\sqrt{n}}\right) - \min_{\theta\in(a,b)} P\left(|Z| > \frac{|\bar{x}-\theta|}{\sigma/\sqrt{n}}\right)}{\max_{\theta\in(a,b)} P\left(|Z| > \frac{|\bar{x}-\theta|}{\sigma/\sqrt{n}}\right) - \min_{\theta\in(a,b)} P\left(|Z| > \frac{|\bar{x}-\theta|}{\sigma/\sqrt{n}}\right)}$$

New Bayesian Evidence

From the Bayesian viewpoint, we propose a new measure of evidence

$$B(X) = P(|\theta - \hat{\theta}| < |\theta_0 - \hat{\theta}| \mid X), \tag{3}$$

where $\hat{\theta}$ is the point estimate of θ and $P(\theta \mid X)$ is the posterior distribution of θ given X.

- Note that $B(X) \in [0,\ 1]$.
- The large value of $B(X)$ indicates that there is a large distance between θ_0 and $\hat{\theta}$ in the probability viewpoint.

Posterior distribution

The posterior distribution of θ (i.e., $P(\theta \mid X)$) is given by

$$P(\theta \mid X) \propto L(\theta, X)\pi(\theta), \tag{4}$$

where $L(\theta, X)$ is the likelihood function and $\pi(\theta)$ is the prior density of θ.

- When σ^2 is known, since $\theta \in (a,\ b)$, we choose $\pi(\theta) = 1/(b - a)$.
- When σ^2 is unknown, we choose $\pi(\theta) = 1(/(b-1)\sigma^2)$, which is often called noninformative prior in Bayesian statistics.

Theorems

Theorem (1)

If the parameter space is unbounded, then the value of $B(\theta)$ is equal to the usual p-value of UMPU test, that is, $B(X) = p(\bar{x})$ if $a = -\infty$ and $b = \infty$.

Theorem (2)

The proposed measure of evidence is also a UMPU test.

Data Set Description

- Consider a real data consisting of 189 infant birth weights collected at Baystate Medical Center, Springfield, Massachusetts, in 1986.
- The population is the entire data set, 189 infant birth weights.
- The mean and variance of this dataset are 2944 g and 729^2 g.
- Assume that infant birth weight follows a normal distribution $N(\theta, \sigma^2)$.
- We are interested in testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_a : \theta \neq \theta_0. \tag{5}$$

Example 1

- Suppose that the bounded values are $a = 2700$ and $b = 3200$
- We have a sample of five weights:

$$2600, \quad 2055, \quad 3062, \quad 3232, \quad 1970.$$

- The mean of the sample is 2583.8.
- We assume that σ^2 is known and equal to $\sigma^2 = 729^2$.
- We are interested in testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_a : \theta \neq \theta_0. \tag{6}$$

Example 1

- Consider $\theta_0 = 2944$. The usual p-value of the UMPU test is

$$p(\bar{x}) = P\left(|Z| > \frac{|2944 - 2583.8|}{729/\sqrt{5}}\right) = 0.269,$$

and if the choose the point estimate $\hat{\theta} = \bar{x}$, then

$$B(X) = P(|\theta - \hat{\theta}| < |2944 - 2583.8| \mid X) = 0.7233. \quad (7)$$

- In this case, when H_0 is true, both measures of evidence fail to reject H_0 when the significance level is $\alpha = 0.1$.

Example 1

- Consider $\theta_0 = 3100$. The usual p-value of the UMPU test is

$$p(\bar{x}) = P\left(|Z| > \frac{|3100 - 2583.8|}{729/\sqrt{5}}\right) = 0.113,$$

and if the choose the point estimate $\hat{\theta} = \bar{x}$, then

$$B(X) = P(|\theta - \hat{\theta}| < |2944 - 2583.8| \mid X) = 0.0823. \quad (8)$$

- In this case, when H_0 is not true, the p-value fails to reject H_0 when the significance level is $\alpha = 0.1$.
- The proposed measure of evidence $B(X)$ is more appropriate.

Example 2

- We conduct a simulation study to compare the performances of the p-value and the proposed evidence as well as the mortified p-value $r(\bar{x})$.
- Choose a sample with size n for the population of 198 infant weights, then calculate the value of each evidence.
- Replicate the process 1000 times.
- Calculate the proportion of each value of the method less than $\alpha = 0.05$.

Comparison of different procedures

We are interested in testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_a : \theta \neq \theta_0.$$

θ_0	H_0 is True or False	n	p-value ≤ 0.05	$r(\bar{x}) \leq 0.05$	B(X) ≤ 0.05
2700	false	10	0.173	0.479	0.590
2800	false	10	0.081	0.102	0.102
2944	true	10	0.049	0.003	0.003
3000	false	10	0.053	0.009	0.009
3100	false	10	0.070	0.132	0.132
3200	false	10	0.151	0.558	0.648

Table: The Proportion of the p-values and the modified p-values as well as the proposed measure of evidence less than 0.05 based on 1000 replicates. The true value of $\theta = 2944$, and the parameter space is $(2700, 3200)$.

Comparison of different procedures

We are interested in testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_a : \theta \neq \theta_0.$$

θ_0	H_0 is True or False	n	p-value ≤ 0.05	$r(\bar{x}) \leq 0.05$	B(X) ≤ 0.05
2700	false	20	0.294	0.511	0.637
2800	false	20	0.135	0.152	0.210
2944	true	20	0.041	0.019	0.019
3000	false	20	0.062	0.050	0.050
3100	false	20	0.105	0.166	0.237
3200	false	20	0.301	0.557	0.687

Table: The Proportion of the p-values and the modified p-values as well as the proposed measure of evidence less than 0.05 based on 1000 replicates. The true value of $\theta = 2944$, and the parameter space is $(2700, 3200)$.

Comparison of different procedures

We are interested in testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_a : \theta \neq \theta_0.$$

θ_0	H_0 is True or False	n	p-value ≤ 0.05	$r(\bar{x}) \leq 0.05$	B(X) ≤ 0.05
2600	false	30	0.720	0.819	0.917
2700	false	30	0.442	0.450	0.578
2800	false	30	0.155	0.174	0.157
2944	true	30	0.042	0.019	0.032
3000	false	30	0.042	0.054	0.107
3100	false	30	0.167	0.292	0.471

Table: The Proportion of the p-values and the modified p-values as well as the proposed measure of evidence less than 0.05 based on 1000 replicates. The true value of $\theta = 2944$, and the parameter space is $(2600, 3100)$.

Conclusion

- The performance of the proposed evidence is better than the usual p-value when the parameter space is bounded.
- In most situations, the performance of the proposed evidence is also better than the modified p-values.
- Similar results are obtained for the case where σ^2 is unknown.
- It is also of interest to compare the confidence intervals based on these procedures.s