Hindawi Publishing Corporation
Journal of Probability and Statistics
Volume 2010, Article ID 139856, 9 pages
doi:10.1155/2010/139856

Research Article

A Note on Confidence Interval for the Power of the One Sample t Test

A. Wong

Department of Mathematics and Statistics, York University, Toronto, Ontario, Canada M3J 1P3

Correspondence should be addressed to A. Wong, august@yorku.ca

Received 26 January 2010; Revised 22 June 2010; Accepted 12 August 2010

Academic Editor: Junbin B. Gao

In introductory statistics texts, the power of the test of a one-sample mean when the variance is known is widely discussed. However, when the variance is unknown, the power of the Student's t-test is seldom mentioned. In this note, a general methodology for obtaining inference concerning a scalar parameter of interest of any exponential family model is proposed. The method is then applied to the one-sample mean problem with unknown variance to obtain a $(1-\gamma)100\%$ confidence interval for the power of the Student's t-test that detects the difference $(\mu-\mu_0)$. The calculations require only the density and the cumulative distribution functions of the standard normal distribution. In addition, the methodology presented can also be applied to determine the required sample size when the effect size and the power of a size α test of mean are given.

1. Introduction

Let $(x_1,\ldots,x_n)$ be a random sample from a normal distribution with mean μ and variance σ^2. As presented in any introductory statistics text, such as Mandenhall et al. [1, page 425], a $(1-\gamma)100\%$ confidence interval for σ^2 is

$$(L_{\sigma^2}, U_{\sigma^2}) = \left(\frac{(n-1)s^2}{\chi^2_{n-1,1-\gamma/2}}, \frac{(n-1)s^2}{\chi^2_{n-1,\gamma/2}}\right), \tag{1.1}$$

where $\overline{x} = \sum x_i/n$, $s^2 = \sum (x_i-\overline{x})^2/(n-1)$, and $\chi^2_{\nu,\delta}$ is the 100δth percentile of the χ^2 distribution with ν degrees of freedom. Moreover, for testing

$$H_0 : \mu = \mu_0 \quad \text{versus } H_a : \mu = \mu_0 + k\sigma \quad k > 0, \tag{1.2}$$

the null hypothesis will be rejected at significance level α if

$$\frac{\overline{x} - \mu_0}{s/\sqrt{n}} > t_{n-1,1-\alpha}, \tag{1.3}$$

where $t_{\nu,\delta}$ is the 100δth percentile of the Student's t distribution with ν degrees of freedom. Although the power of this test is rarely discussed in introductory statistics texts, Lehmann [2] proved that the probability of committing Type II error of a size α test with the hypotheses stated in (1.2) is

$$\beta = G_{n-1,k\sqrt{n}}(t_{n-1,1-\alpha}), \tag{1.4}$$

where $k = (\mu - \mu_0)/\sigma$ is the effect size and $G_{\nu,\lambda}(\cdot)$ is the cumulative distribution function of the noncentral t distribution with ν degrees of freedom and noncentrality λ. Note that the calculation of β involves the unknown σ. A naive point estimate of β is

$$\widehat{\beta} = G_{n-1,\widehat{k}\sqrt{n}}(t_{n-1,1-\alpha}), \tag{1.5}$$

where $\widehat{k} = (\overline{x} - \mu_0)/s$. Thus, the corresponding point estimate of the power of the size α test that detects the difference $(\mu - \mu_0)$ is $1 - \widehat{\beta}$.

In Section 2, a general methodology is proposed for obtaining inference concerning a scalar parameter of interest of an exponential family model. Applying the general methodology to the one-sample mean problem with unknown variance, a $(1 - \gamma)100\%$ confidence interval for $1 - \beta$ is derived. This interval estimate will depend only on the evaluation of the density and the cumulative distribution functions of the standard normal distribution. The methodology can also be used to determine the required sample size when the effect size and the power of a size α test are fixed. Numerical examples are presented in Section 3 to illustrate the accuracy of the proposed method. Finally, some concluding remarks are given in Section 4.

2. Confidence Interval for the Power of the Test and Sample Size Calculation

From (1.1), for a given $(\mu - \mu_0)$ value, a $(1 - \gamma)100\%$ confidence interval for $(\mu - \mu_0)\sqrt{n}/\sigma$ is

$$\left(\frac{(\mu - \mu_0)\sqrt{n}}{\sqrt{U_{\sigma^2}}}, \frac{(\mu - \mu_0)\sqrt{n}}{\sqrt{L_{\sigma^2}}} \right). \tag{2.1}$$

Hence, from (1.4), the corresponding confidence interval for β is

$$(L_\beta, U_\beta) = \left(G_{n-1,(\mu-\mu_0)\sqrt{n}/\sqrt{U_{\sigma^2}}}(t_{n-1,1-\alpha}), G_{n-1,(\mu-\mu_0)\sqrt{n}/\sqrt{L_{\sigma^2}}}(t_{n-1,1-\alpha}) \right). \tag{2.2}$$

Finally, a $(1-\gamma)100\%$ confidence interval for the power of a size α test that detects the difference $(\mu-\mu_0)$ is

$$\left(1-U_\beta, 1-L_\beta\right). \tag{2.3}$$

Evaluating (2.3) requires the cumulative distribution function of the noncentral t distribution, which is generally not discussed in introductory statistics texts. In statistics literature, various approximations of $G_{\nu,\lambda}(\cdot)$ have been proposed. For the rest of this section, a simple and accurate approximation of $G_{\nu,\lambda}(\cdot)$ will be derived.

Let $X_1,\ldots,X_n$ be identically independently normally distributed random variables with mean μ and variance σ^2. It is well known that $\overline{X}=\sum X_i/n$ and $(n-1)S^2/\sigma^2=\sum(X_i-\overline{X})^2/\sigma^2$ are independently distributed as normal with mean μ and variance σ^2/n and χ^2 with $(n-1)$ degrees of freedom, respectively. Let

$$T^*=\frac{\sqrt{n}\left[\left(\overline{X}-\mu\right)/\sigma\right]-\sqrt{n}z_p}{S/\sigma}=\frac{\sqrt{n}}{S}\left(\overline{X}-\mu-z_p\sigma\right), \tag{2.4}$$

where z_p denotes the 100pth percentile of the standard normal distribution, then T^* follows a noncentral t distribution with $(n-1)$ degrees of freedom and noncentrality $-\sqrt{n}z_p$.

Now, consider a sample $(x_1,\ldots,x_n)$ from a normal distribution with mean μ and variance σ^2. Let the parameter of interest be

$$\psi=\frac{\sqrt{n}}{s}\left(\overline{x}-\mu-z_p\sigma\right), \tag{2.5}$$

where $\overline{x}=\sum x_i/n$ and $s^2=\sum(x_i-\overline{x})^2/(n-1)$, then the log-likelihood function can be written as

$$\ell(\theta)=\ell\left(\psi,\sigma^2\right)=-\frac{n}{2}\log\sigma^2-\frac{1}{2\sigma^2}s^2\left(n-1+\psi^2\right)+\delta s\frac{\psi}{\sqrt{\sigma^2}}, \tag{2.6}$$

where $\delta=-\sqrt{n}z_p$. Denote that

$$p(\psi)=P\left(T^*\le\psi\right)=G_{t_{n-1,\delta}}(\psi). \tag{2.7}$$

The overall maximum likelihood estimate (MLE) of θ, $\widehat{\theta}=(\widehat{\psi},\widehat{\sigma}^2)'=(\delta\sqrt{(n-1)/n}, ((n-1)/n)s^2)'$ is obtained by solving $(\partial\ell(\theta)/\partial\theta)|_{\theta=\widehat{\theta}}=0$, and the determinant of the observed information matrix evaluated at the overall mle is

$$\left|j_{\theta\theta'}\left(\widehat{\theta}\right)\right|=\left|\frac{\partial^2\ell(\theta)}{\partial\theta\partial\theta'}\right|_{\theta=\widehat{\theta}}=\frac{n^4}{2(n-1)^3s^4}. \tag{2.8}$$

The constrained mle of θ at a fixed ψ, $\widehat{\theta}_\psi = (\psi, \widehat{\sigma}^2_\psi) = (\psi, s^2A^2/4n^2)$, where

$$A = \sqrt{\delta^2\psi^2 + 4n(n - 1 + \psi^2)} - \delta\psi, \tag{2.9}$$

is obtained by solving $(\partial\ell(\theta)/\partial\sigma^2)|_{\theta=\widehat{\theta}_\psi} = 0$. Moreover, the determinant of the observed nuisance information matrix evaluated at the constrained mle is

$$\left|j_{\sigma^2\sigma^2}\left(\widehat{\theta}_\psi\right)\right| = \left|\frac{\partial^2\ell(\theta)}{\partial\sigma^2\partial\sigma^2}\right|_{\theta=\widehat{\theta}_\psi} = \frac{8n^5}{s^4A^5}(A + \delta\psi). \tag{2.10}$$

Hence, the signed log-likelihood ratio statistic is

$$\begin{aligned} r = r(\psi) &= \operatorname{sgn}(\widehat{\varphi} - \psi)\left\{2[\ell(\widehat{\theta}) - \ell(\widehat{\theta}_\psi)]\right\}^{1/2} \\ &= \operatorname{sgn}(\widehat{\varphi} - \psi)\left\{-n\log\left[\frac{4n(n-1)}{A^2}\right] + \delta^2 - \frac{2n\delta\psi}{A}\right\}^{1/2}. \end{aligned} \tag{2.11}$$

It is well known that r is asymptotically distributed as the standard normal distribution with rate of convergence $O(n^{-1/2})$. Hence, $p(\psi)$ can be approximated by $\Phi(r)$ where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. It is important to note that r is reparameterization invariant.

In statistics literatures, various likelihood-based small sample asymptotic methods have been proposed. In particular, if the model is a canonical exponential family model and the canonical parameter is $\theta = (\psi, \lambda')'$, Lugannani and Rice [3] derive

$$p(\psi) = 1 - \Phi(r) - \phi(r)\left\{\frac{1}{r} - \frac{1}{q}\right\}, \tag{2.12}$$

where $\phi(\cdot)$ is the density function of the standard normal distribution, r is defined in (2.11), and q takes the form

$$q = q(\psi) = (\widehat{\varphi} - \psi)\left\{\frac{\left|j_{\theta\theta'}\left(\widehat{\theta}\right)\right|}{\left|j_{\lambda\lambda'}\left(\widehat{\theta}_\psi\right)\right|}\right\}^{1/2}. \tag{2.13}$$

This approximation has a rate of convergence $O(n^{-3/2})$. It is important to note that r is reparameterization invariant whereas q is not.

For a general exponential family model with canonical parameter $\varphi = \varphi(\theta)$ and a scalar parameter $\psi = \psi(\theta)$, to obtain inference concerning ψ based on the Lugannani and Rice (1980) [3] method, r remains unchanged as in (2.11) because it is reparameterization invariant, but q has to be re-expressed in the canonical parameter scale, ϕ scale. To achieve this, let $\varphi_\theta(\theta)$ and $\varphi_\lambda(\theta)$ be the derivatives of $\varphi(\theta)$ with respect to θ and λ, respectively. Denote $\varphi^{\psi}(\theta)$ to be the row of $\varphi_{\theta}^{-1}(\theta)$ that corresponds to ψ, and $\|\varphi^{\psi}(\theta)\|^2$ is the square length of the vector $\varphi^{\psi}(\theta)$. Let $\chi(\theta)$ be a rotated coordinate of $\varphi(\theta)$ that agrees with $\psi(\theta)$ at $\widehat{\theta}_{\psi}$. Then

$$\chi(\theta) = \frac{\varphi^{\psi}\left(\widehat{\theta}_{\psi}\right)}{\left\|\varphi^{\psi}\left(\widehat{\theta}_{\psi}\right)\right\|}\varphi(\theta) \tag{2.14}$$

can be viewed operationally as the scalar parameter of interest in $\varphi(\theta)$ scale.

Since $\ell(\theta) = \ell(\varphi(\theta))$, by the chain rule in differentiation, we have

$$\left|j_{\varphi\varphi'}\left(\widehat{\theta}\right)\right| = \left|j_{\theta\theta'}\left(\widehat{\theta}\right)\right|\left|\varphi_{\theta}\left(\widehat{\theta}\right)\right|^{-2}, \qquad \left|j_{(\lambda\lambda')}\left(\widehat{\theta}_{\psi}\right)\right| = \left|j_{\lambda\lambda'}\left(\widehat{\theta}_{\psi}\right)\right|\left|\varphi'_{\lambda}\left(\widehat{\theta}_{\psi}\right)\varphi_{\lambda}\left(\widehat{\theta}_{\psi}\right)\right|^{-1}. \tag{2.15}$$

Hence, an estimated variance for $|\chi(\widehat{\theta}) - \chi(\widehat{\theta}_{\psi})|$ in $\varphi(\theta)$ scale is $|j_{(\lambda\lambda')}(\widehat{\theta}_{\psi})|/|j_{\varphi\varphi'}(\widehat{\theta})|$. Thus, $q = q(\psi)$, as defined in (2.13) and expressed in $\varphi(\theta)$ scale, is

$$q = q(\psi) = \operatorname{sgn}(\widehat{\psi} - \psi)\left|\chi\left(\widehat{\theta}\right) - \chi\left(\widehat{\theta}_{\psi}\right)\right|\left\{\frac{\left|j_{\varphi\varphi'}\left(\widehat{\theta}\right)\right|}{\left|j_{(\lambda\lambda')}\left(\widehat{\theta}_{\psi}\right)\right|}\right\}^{1/2}. \tag{2.16}$$

Therefore, $p(\psi)$ can be obtained from (2.12) with r and q being defined in (2.11) and (2.17), respectively.

Note that the model being considered is an exponential family model with canonical parameter

$$\varphi(\theta) = \left(\frac{1}{\sigma^2}, \frac{1}{\sigma^2}\left(\overline{x} - \frac{s}{\sqrt{n}}\psi - z_p\sqrt{\sigma^2}\right)\right)'. \tag{2.17}$$

From (2.17), we have

$$\varphi_{\theta}(\theta) = \begin{pmatrix} 0 & -\dfrac{1}{\sigma^4} \\ -\dfrac{s}{\sqrt{n}\sigma^2} & -\dfrac{\overline{x}}{\sigma^4} + \dfrac{s\psi}{\sqrt{n}\sigma^4} - \dfrac{\delta}{2\sqrt{n}(\sigma^2)^{3/2}} \end{pmatrix},$$
$$\left|\varphi_{\theta}\left(\widehat{\theta}\right)\right| = -\frac{s}{\sqrt{n}\widehat{\sigma}^6}, \qquad \left|\varphi'_{\sigma^2}\left(\widehat{\theta}_{\psi}\right)\varphi_{\sigma^2}\left(\widehat{\theta}_{\psi}\right)\right| = \frac{1}{\widehat{\sigma}_{\psi}^8}\left(1 + B^2\right), \tag{2.18}$$

where

$$B = -\overline{x} + \frac{s\psi}{\sqrt{n}} - \frac{\delta\sqrt{\widehat{\sigma}_{\psi}^2}}{2\sqrt{n}}. \tag{2.19}$$

Moreover, by obtaining the inverse of $\varphi_\theta(\theta)$, we have

$$\varphi^{\psi}(\theta) = \frac{\sqrt{n}\sigma^2}{s}(-B, -1). \tag{2.20}$$

Hence, from (2.14), we can obtain

$$\left|\chi\left(\widehat{\theta}\right) - \chi\left(\widehat{\theta}_{\psi}\right)\right| = \frac{1}{\sqrt{1+B^2}}\left|-\frac{\sqrt{n}\psi}{(n-1)s} + \frac{\delta A^2 + 4\delta n(n-1)}{4\sqrt{n}(n-1)sA}\right|. \tag{2.21}$$

Thus, from (2.16), we have

$$q = q(\psi) = \operatorname{sgn}(\widehat{\varphi} - \psi)\frac{\left|\delta A^2 - 4n\psi A + 4\delta n(n-1)\right|}{\sqrt{A+\delta\psi}}\frac{\sqrt{n(n-1)}}{A^{5/2}}. \tag{2.22}$$

Finally, $p(\psi) = G_{n-1,\delta}(\psi)$ can be approximated from (2.12) with rate of convergence $O(n^{-3/2})$.

By reindexing all the necessary equations, we have

$$G_{n-1,k\sqrt{n}}(t_{n-1,1-\alpha}) = 1 - \Phi(r) - \phi(r)\left\{\frac{1}{r} - \frac{1}{q}\right\}, \tag{2.23}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and cumulative distribution functions of the standard normal distribution, and

$$\begin{aligned}
r &= \operatorname{sgn}\left(k\sqrt{n-1} - t_{n-1,1-\alpha}\right)\left\{-n\log\frac{4n(n-1)}{A^2} + nk^2 - \frac{2kn^{3/2}t_{n-1,1-\alpha}}{A}\right\}^{1/2}, \\
q &= \operatorname{sgn}\left(k\sqrt{n-1} - t_{n-1,1-\alpha}\right)\frac{n(n-1)}{A^{5/2}}\frac{\left|k\sqrt{n}A^2 - 4nAt_{n-1,1-\alpha} + 4kn^{3/2}(n-1)\right|}{\sqrt{A + kn^{1/2}t_{n-1,1-\alpha}}}, \\
A &= -k\sqrt{n}t_{n-1,1-\alpha}. + \sqrt{k^2nt^2_{n-1,1-\alpha} + 4n(n-1+t_{n-1,1-\alpha})}.
\end{aligned} \tag{2.24}$$

Finally, with a predetermined effect size k and power of a size α test, the sample size can be obtained by iterations.

Note that DiCiccio and Martin [4] derived an asymptotic approximation of marginal tail probabilities for a real-valued function of a random vector where the function has continuous gradient that does not vanish at the mode of the joint density of the random vector. Applied to the noncentral t distribution problem, the results are identical. Nevertheless, the approach of DiCiccio and Martin [4] is quite different from the proposed method. More specifically, DiCiccio and Martin [4] worked directly from the log density and treated the parameters as fixed whereas the proposed method works from the log-likelihood function where the data are observed.

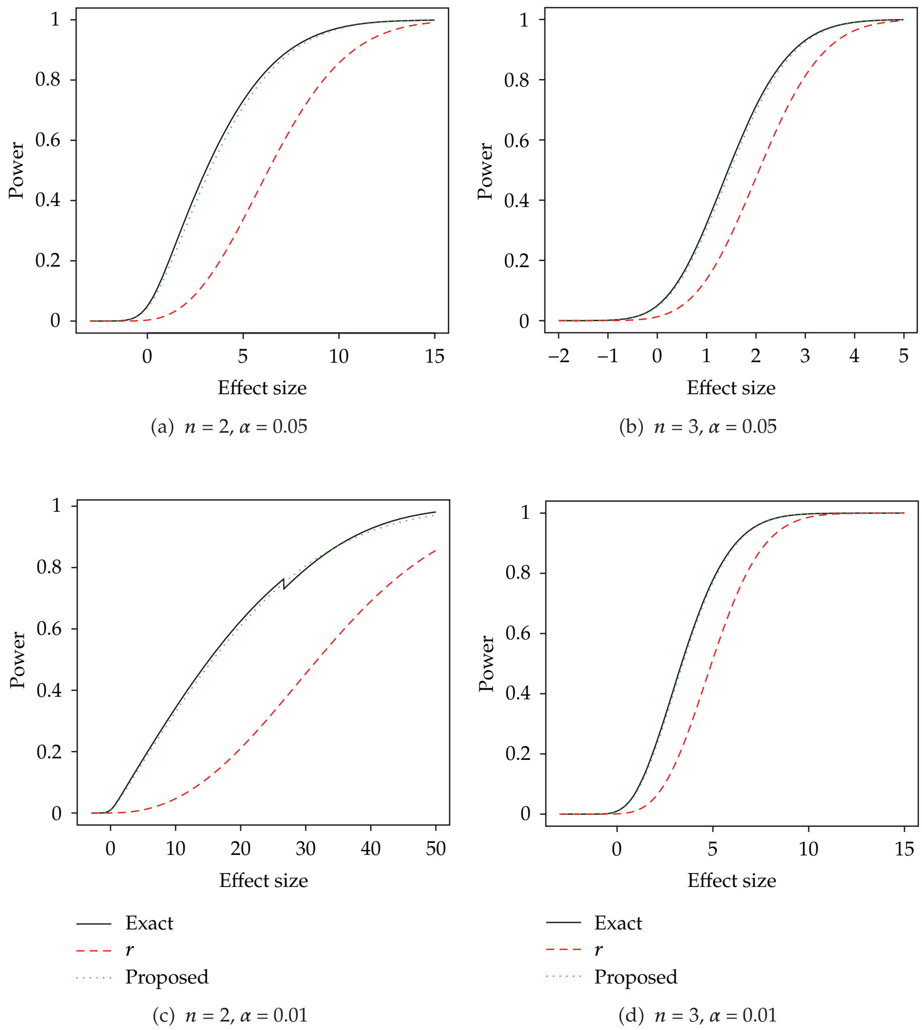

Figure 1

3. Numerical Example

Figure 1 plots the power function of a one-sample t test against the effect size k for $n = 2, 3$ and $\alpha = 0.05, 0.01$. The exact method is obtained from the built-in cumulative distribution function of the noncentral t distribution in R. From the plot, it is clear that the signed log-likelihood ratio does not provide satisfactory results. The proposed method and the built-in function of R are very close even when the sample size is 2. It is interesting to note that the built-in function of R has a discontinuity point in the $n = 2$, $\alpha = 0.01$ case.

Now, consider the data set recorded in Mandenhall et al. [1, page 103]

$$0.46, \quad 0.61, \quad 0.52, \quad 0.48, \quad 0.57, \quad 0.54. \tag{3.1}$$

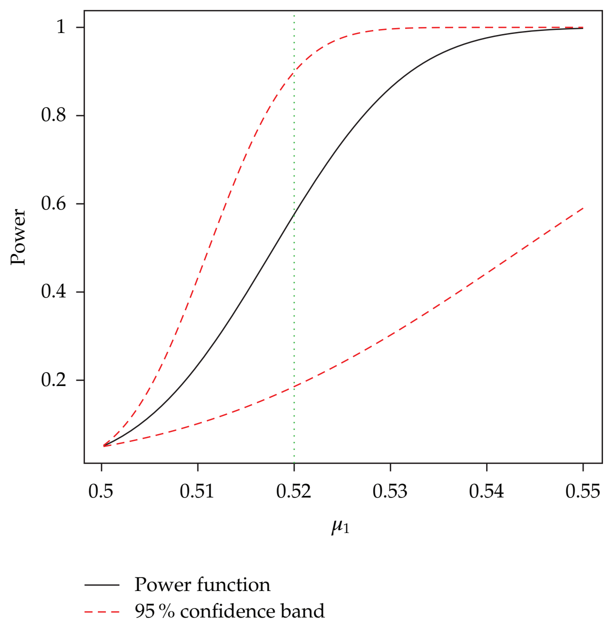

Figure 2: Power function and the 95% confidence band.

For testing the hypothesis

$$H_0 : \mu = 0.5 \quad \text{versus } H_a : \mu = \mu_1 > 0.5, \tag{3.2}$$

the power function of a size 0.05 test and the corresponding 95% confidence bands are plotted in Figure 2. From Figure 2, the approximated power at $\mu_1 = 0.52$ is 0.5764. Furthermore, the 95% confidence interval for the power of the above test when $\mu_1 = 0.52$ is $(0.1856,\ 0.8992)$. At first, the confidence interval seems too wide. However, by examining (2.3), the result is not too surprising because (2.3) depends on (1.1). Since χ^2 distribution is a skewed distribution, by defining the confidence interval of σ^2 to have equal tail coverage, (1.1) is a wide interval and hence (2.3) is a wide interval.

Finally, to illustrate the determination of the sample size, let the effect size be 0.8, and at $\alpha = 0.025$, let the power be at least 0.9, then the proposed method gives $n = 19$ with power 0.909.

4. Summary and Conclusion

The $(1-\gamma)100\%$ confidence interval for the power of the size α Student's t-test detecting the difference $(\mu-\mu_0)$ is presented. The major advantages of the presented confidence interval are that it depends only on the evaluations of the density and cumulative distribution functions of the standard normal distribution and that it is extremely accurate. The R source code is available from the author upon request.

As a final note, the proposed method can be applied to any distribution that belongs to the exponential family model with known canonical parameters. Although the method depends on the correct specification of the underlying distribution, Fraser et al. [5] examined a special case when the error distribution of the regression model is misspecified and the likelihood-based method still gives results that are more accurate than the existing Central Limit Theorem-based approximations.

References

[1] W. Mandenhall, R. Beaver, B. Beaver, and S. Ahmed, *Introduction to Probability and Statistics*, Nelson, Torrent, Canada, 2009.

[2] E. L. Lehmann, *Testing Statistical Hypotheses*, John Wiley & Sons, New York, NY, USA, 1959.

[3] R. Lugannani and S. Rice, "Saddle point approximation for the distribution of the sum of independent random variables," *Advances in Applied Probability*, vol. 12, no. 2, pp. 475–490, 1980.

[4] T. J. DiCiccio and M. A. Martin, "Approximations of marginal tail probabilities for a class of smooth functions with applications to Bayesian and conditional inference," *Biometrika*, vol. 78, no. 4, pp. 891–902, 1991.

[5] D. A. S. Fraser, A. Wong, and J. Wu, "Regression analysis, normal and nonnormal: accurate P-values from likelihood analysis," *Journal of the American Statistical Association*, vol. 94, no. 448, pp. 1286–1295, 1999.