ON CERTAIN LIKELIHOOD-RATIO TESTS ASSOCIATED WITH THE EXPONENTIAL DISTRIBUTION

By Edward Paulson

Washington, D.C.

Various likelihood-ratio tests and their distributions in samples from a population having the elementary probability law $\frac{1}{\sigma} e^{-(x-B)/\sigma}$, $B \leq x \leq \infty$, have been studied by Neyman and Pearson [1] and Sukhatme [2]. In this note the power functions and the question of bias of several likelihood-ratio tests will be investigated. The exponential distribution appears to be appropriate for dealing with problems involving the intervals of time between events which tend to be random, as for example the interval between consecutive telephone calls, or the interval between consecutive accidents to the same worker.

To test the hypothesis H' that the location parameter B is equal to some fixed value, it being assumed that the scale parameter σ is known, we can for simplicity take the set Ω of admissible populations from which the sample might have been drawn to be $\{-\infty < B < +\infty, \sigma = 1\}$, while the subset ω from which the sample must come when the hypothesis is true is $\{B = 0, \sigma = 1\}$. Then the likelihood-ratio λ_1 for testing this hypothesis is

$$\lambda_1 = \frac{P(\omega \text{ max.})}{P(\Omega \text{ max.})} = \frac{e^{-\sum_{i=1}^{n} x_i}}{e^{-\sum_{i=1}^{n} (x_i - x_1)}} = e^{-nx_1},$$

where x_1 is the smallest observation in a random sample of n. The region of acceptance of this hypothesis consists of all points in sample space for which

$$\lambda_{1\epsilon} \leq \lambda_1 \leq 1,$$

where $\lambda_{1\epsilon}$ is chosen so that $\int_{\lambda_{1\epsilon}}^{1} g_1(\lambda_1)\, d\lambda_1 = 1 - \alpha$, α being the level of significance used and $g\ (\lambda_1)\, d\lambda_1$ being the distribution of λ_1 when B is really equal to zero. The region $\lambda_{1\epsilon} \leq \lambda_1 \leq 1$ is equivalent to the region in the sample space for which

$$0 \leq x_1 \leq k_1\, ;\, k_1 = -\frac{\log \lambda_{1\epsilon}}{n}.$$

For any value of B the distribution of x_1 is known [3] to be

$$\phi_1(x_1)\, dx_1 = ne^{-n(x_1 - B)}\, dx_1 .$$

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

Setting $B = 0$, the relationship between k_1 and α is

$$\int_0^{k_1} ne^{-nx_1}\,dx_1 = 1 - \alpha, \qquad \text{so} \qquad e^{-nk_1} = \alpha.$$

When $B \leq 0$, the power function $P(B)$, for this test is

$$P(B) = 1 - \int_0^{k_1} ne^{-n(x_1 - B)}\,dx_1 = 1 - e^{nB}[1 - \alpha].$$

When $0 \leq B \leq k_1$, $P(B) = 1 - \int_B^{k_1} ne^{-n(x_1-B)}\,dx_1 = \alpha e^{nB}$. When $B \geq k_1$, $P(B) = 1$.

Since $e^{nB} > 1$ if $B > 0$ and also $e^{nB} < 1$ if $B < 0$, $P(B)$ is obviously $> \alpha$ if $B \neq 0$. This test is therefore completely unbiased in the sense of Daly [4]. In addition, it is not difficult to prove that this test has the unusual property of being a uniformly most powerful test with respect to all alternatives.

To test the hypothesis H'' that the location parameter is equal to some fixed value, say $B = 0$, when the scale parameter σ is unknown, the likelihood-ratio is easily seen to be

$$\lambda_2 = \left[\frac{\sum_{i=1}^{n}(x_i - x_1)}{\sum_{i=1}^{n} x_i}\right]^n = \left[\frac{1}{1 + \dfrac{nx_1}{\sum_{i=1}^{n}(x_i - x_1)}}\right]^n.$$

The region of acceptance consists of all points in the sample space for which $\lambda_{2\epsilon} \leq \lambda_2 \leq 1$ where $\int_{\lambda_{2\epsilon}}^{1} g_2(\lambda_2)\,d\lambda_2 = 1 - \alpha$. This is equivalent to the region

$$(1) \qquad 0 \leq \left[\frac{n(n-1)x_1}{\sum_{i=1}^{n}(x_i - x_1)} = t\right] \leq k_2; \qquad k_2 = (n-1)\,\frac{(1 - \lambda_{2\epsilon}^{1/n})}{\lambda_{2\epsilon}^{1/n}}.$$

The relation between k_2 and α is easily found from the distribution of t when $B = 0$, which is known to be [3]

$$\phi_2(t)\,dt = \frac{dt}{\left[1 + \dfrac{t}{n-1}\right]^n}.$$

Therefore $\int_0^{k_2} \phi_2(t)\,dt = 1 - \alpha$, so $\left[1 + \dfrac{k_2}{n-1}\right]^{-(n-1)} = \alpha$.

It is somewhat easier to find the power function of this test by considering the region of acceptance as made up of points in the x_1, s plane for which

$$0 \leq x_1 \leq \frac{k_2 s}{n} \quad \text{where } s = \frac{\sum_{i=1}^{n}(x_i - x_1)}{n-1},$$

which is identical with the region in (1).

The joint distribution of x_1 and s is [3]

$$\psi_1(x_1, s)\, dx_1\, ds = \phi_3(x_1)\, dx_1 \cdot \phi_4(s)\, ds,$$

where

$$\phi_3(x_1)\, dx_1 = \frac{n}{\sigma} e^{-n(x_1-B)/\sigma}\, dx_1$$

and

$$\phi_4(s)\, ds = \frac{\left(\frac{n-1}{\sigma}\right)^{n-1} s^{n-2} e^{-(n-1)s/\sigma}\, ds}{(n-2)!}.$$

When $B \leq 0$, the power function $P(B)$ of this test is

$$P(B) = 1 - \int_0^\infty ds \int_0^{k_2 s/n} \psi_1(x_1 s)\, dx_1 = 1 - e^{nB/\sigma}[1 - \alpha].$$

When $B \geq 0$, the power function is

$$\begin{aligned} P(B) &= 1 - \int_{Bn/k_2}^\infty ds \int_B^{k_2 s/n} \psi_1(x_1, s)\, dx_1 \\ &= \alpha e^{nB/\sigma} + I\left[n-1; \frac{n(n-1)B}{\sigma k_2}\right] - \alpha e^{nB/\sigma} I\left[n-1; \frac{n(n-1+k_2)B}{\sigma k_2}\right], \end{aligned} \tag{2}$$

where $\mathrm{I}[p; x] = \dfrac{\Gamma_x(p)}{\Gamma(p)} = \dfrac{\int_0^x x^{p-1} e^{-x}\, dx}{\int_0^\infty x^{p-1} e^{-x}\, dx}$,

which is the form in which the Incomplete Gamma Function has been tabulated [5].

Since σ must be positive, $e^{nB/\sigma} < 1$ if $B < 0$ and therefore $P(B) > \alpha$ in the interval $-\infty < B < 0$. To show that $P(B)$ is $> \alpha$ in the interval $0 < B < \infty$, it is simpler to work with the expression for $P(B)$ as a double integral in (2), than to differentiate the power function directly. Performing the integration with respect to x_1,

$$P(B) = 1 + \int_{Bn/k_2}^\infty [e^{-(k_2 s - Bn)/\sigma} - 1] \cdot \phi_4(s)\, ds.$$

Differentiating with respect to B,

$$P'(B) = \int_{Bn/k_2}^\infty \frac{n}{\sigma} e^{-(k_2 s - Bn)/\sigma} \phi_4(s)\, ds.$$

The integral expression for $P'(B)$ is obviously positive. Therefore since for $B > 0$ the derivative is always positive the function must be monotomically increasing in this interval ($0 < B < +\infty$), so $P(B)$ is $> \alpha$ when $B > 0$. Therefore this test is also completely unbiased.

We now consider the hypothesis H''' that two samples are drawn from exponential distributions with the same location parameter, assuming it is known the samples must have come from two exponential distributions with the same scale parameter. Given a sample of n_1 values of x drawn from $\frac{1}{\sigma} e^{-(x-B_1)/\sigma}\, dx$ and another independent sample of n_2 values of y drawn from $\frac{1}{\sigma} e^{-(y-B_2)/\sigma}\, dy$, the hypothesis we wish to test is that $B_2 = B_1$. Let x_1 be the smallest of the n_1 values of x and y_1 be the smallest of the n_2 values of y, let L be the smallest of the $n_1 + n_2 = N$ values of both x and y. Then the likelihood ratio for this hypothesis is

$$\lambda_3 = \left[\frac{\sum_{i=1}^{n} (x_i - x_1) + \sum_{i=1}^{n_2} (y_i - y_1)}{\sum_{i=1}^{n_1} (x_i - L) + \sum_{i=1}^{n_2} (y_i - L)}\right]^N = \left[\frac{1}{1 + \frac{z}{u}}\right]^N,$$

where

$$\begin{aligned} z &= n_2(y_1 - x_1), \qquad \text{if } y_1 > x_1 \\ &= n_1(x_1 - y_1), \qquad \text{if } x_1 > y_1, \end{aligned}$$

and

$$u = \sum_{i=1}^{n_1} (x_i - x_1) + \sum_{i=1}^{n_2} (y_i - y_1).$$

The region of acceptance, $\lambda_{3\epsilon} \leq \lambda_3 \leq 1$, is equivalent to the region $0 \leq Z \leq K_3 u$, where K_3 is again a function of α, the level of significance, the exact relation being

$$\int_0^{k_3} \frac{(N-2)\, dt}{(1+t)^{N-1}} = 1 - \alpha, \qquad \text{so} \qquad \frac{1}{(1+k_3)^{N-2}} = \alpha.$$

It is known [3] that u is independent of Z, and that its distribution is

$$\phi_5(u)\, du = \frac{u^{N-3} e^{-u/\sigma}\, du}{\sigma^{N-2}(N-3)!}.$$

The distribution of z is somewhat complicated; but it can be derived by observing that the probability that z lies in any infinitesimal interval $z_1 \pm \frac{1}{2}\, dz_1$ is the sum of the probabilities that $n_2(y_1 - x_1)$ and $n_1(x_1 - y_1)$ lie in that interval and by then using standard methods for finding the distribution of the difference of two variates. For the case $G = B_2 - B_1 \geq 0$, the distribution $f(z)$ of z is

$$\text{(3)} \qquad \begin{aligned} f_1(z)\, dz &= \frac{e^{-n_1 G/\sigma}}{(n_1+n_2)\sigma} [n_1 e^{n_1 z/n_2 \sigma} + n_2 e^{-z/\sigma}]\, dz, \qquad 0 \leq z \leq n_2 G, \\ f_2(z)\, dz &= \frac{[n_1 e^{n_2 G/\sigma} + n_2 e^{-n_1 G/\sigma}] e^{-z/\sigma}\, dz}{(n_1+n_2)\sigma}, \qquad n_2 G \leq z \leq \infty. \end{aligned}$$

For the case $G \leq 0$, the distribution of z can be derived from (3) by interchanging n_1 and n_2 and putting $-G$ in place of G.

The power function of this test can now be derived. For the case $G \geq 0$, the power function $P(G)$ is

$$P(G) = 1 - \left\{\int_0^\infty du \int_0^{n_2 G} f_1(z)\phi_5(u)\,dz - \int_0^{n_2 G/k_3} du \int_{k_3 u}^{n_2 G} f_1(z)\phi_5(u)\,dz + \int_{n_2 G/k_3}^\infty du \int_{n_2 G}^{k_3 u} f_2(z)\phi_5(u)\,dz\right\}. \tag{4}$$

Upon integrating out and simplifying, the power function becomes

$$\begin{aligned} P(G) = {} & \alpha\left(\frac{n_2 e^{-n_1 G/\sigma}}{n_1 + n_2}\right) + I\left[N - 2; \frac{n_2 G}{k_3 \sigma}\right] \\ & + \alpha\left(\frac{n_1 e^{n_2 G/\sigma}}{n_1 + n_2}\right)\left\{1 - I\left[N - 2; \frac{n_2 G(1 + k_3)}{k_3 \sigma}\right]\right\} \\ & - \frac{n_2}{n_1 + n_2} e^{-n_1 G/\sigma}\left(\frac{n_2}{n_2 - n_1 k_3}\right)^{N-2} I\left[N - 2; \frac{G(n_2 - n_1 k_3)}{k_3 \sigma}\right]. \end{aligned}$$

The power function when $G \leq 0$ is easily derived from that for $G \geq 0$ by everywhere interchanging n_1 and n_2 and substituting $-G$ for G.

To show that $P(G) > \alpha$ when $G \neq 0$, it is only necessary to show that the derivative $P'(G)$ of the power function is always positive when $G > 0$, and always negative when $G < 0$. It is again considerably simpler to use the expression for $P(G)$ as a double integral. For the case $G > 0$, integrating with respect to z in (4),

$$\begin{aligned} P(G) = {} & 1 - \frac{n_2}{n_1 + n_2}\left[1 - e^{-G(n_1 + n_2)/\sigma}\right] \\ & + \int_0^{n_2 G/k_3} \frac{n_2 e^{-n_1 G/\sigma}}{n_1 + n_2}\left[e^{n_1 z/n_2 \sigma} - e^{-z/\sigma}\right]_{k_3 U}^{n_2 G} \phi_5(u)\,du \\ & - \int_{n_2 G/k_3}^\infty \frac{(n_1 e^{n_2 G/\sigma} + n_2 e^{-n_1 G/\sigma})}{n_1 + n_2} \cdot \left[-e^{-z/\sigma}\right]_{n_2 G}^{k_3 u} \phi_5(u)\,du, \end{aligned}$$

where $[f(x)]_a^b = f(b) - f(a)$. Upon differentiating and simplifying,

$$\begin{aligned} P'(G) = {} & \frac{n_1 n_2}{(n_1 + n_2)\sigma} \int_0^{n_2 G/k_3} e^{-n_1 G/\sigma}\left[e^{n_1 k_3 u/n_2 \sigma} - e^{-k_3 u/\sigma}\right]\phi_5(u)\,du \\ & + \frac{n_1 n_2}{(n_1 + n_2)\sigma} \int_{n_2 G/k_3}^\infty e^{-k_3 u/\sigma}\left[e^{n_2 G/\sigma} - e^{-n_1 G/\sigma}\right]\phi_5(u)\,du. \end{aligned}$$

Both integrals are easily seen to always be positive, so $P'(G)$ is positive when $G > 0$. In the same manner it can be shown that $P'(G)$ is negative when $G < 0$. Therefore this test is also completely unbiased.

The question of investigating the bias of the likelihood-ratio tests for (a) testing the hypothesis that $\sigma = \sigma_0$ when B is known and (b) testing the hypothesis that $\sigma = \sigma_0$, nothing being known about the value of B, are practically identical with the analogous problems for a normal distribution. The results are also the same, for the λ test for (a) is completely unbiased, while that for (b) is biased.

REFERENCES

[1] J. Neyman and E. S. Pearson, "On the use and interpretation of certain test criteria for purposes of statistical inference," *Biometrika*, Vol. 20a (1928), pp. 221–230.

[2] P. V. Sukhatme, "On the analysis of k samples from exponential populations with special reference to the problem of random intervals," *Stat. Res. Memoirs*, Vol. 1, pp. 94–112.

[3] P. V. Sukhatme, "Tests of significance for samples of the χ^2 population with 2 degrees of freedom," *Annals of Eugenics*, Vol. 8 (1937–38), pp. 54–55.

[4] Joseph F. Daly, "On the unbiased character of likelihood-ratio tests for independence in normal systems," *Annals of Math. Stat.*, Vol. 11 (1940), p. 2.

[5] K. Pearson (Editor), *Tables of the Incomplete Gamma Function*, Biometric Laboratory, London, 1922.