ON STATISTICS INDEPENDENT OF A COMPLETE SUFFICIENT STATISTIC

By D. BASU

Indian Statistical Institute, Calcutta

1. Introduction

If $\{P_\theta\}$, $\theta\epsilon\Omega$, be a family of probability measures on an abstract sample space $\mathfrak{S}$ and T be a sufficient statistic for θ then for a statistic T_1 to be stochastically independent of T it is necessary that the probability distribution of T_1 be independent of θ. The condition is also sufficient if T be a boundedly complete sufficient statistic. Certain well-known results of distribution theory follow immediately from the above considerations. For instance, if $x_1, x_2, \ldots, x_n$ are independent $N(\mu, \sigma)$'s then the sample mean $\bar{x}$ and the sample variance s^2 are mutually independent and are jointly independent of any statistic f (real or vector valued) that is independent of change of scale and origin. It is also deduced that if $x_1, x_2, \ldots, x_n$, are independent random variables such that their joint distribution involves an unknown location parameter θ then there can exist a linear boundedly complete sufficient statistic for θ only if the x's are all normal. Similar characterizations for the Gamma distribution also are indicated.

2. Definitions

Let $(\mathfrak{S}, \mathscr{A})$ be an arbitrary measurable space (the sample space) and let $\{P_\theta\}$, $\theta\epsilon\Omega$, be a family of probability measures on $\mathscr{A}$.

Definition 1: Any measurable transformation T of the sample space $(\mathfrak{S}, \mathscr{A})$ onto a measurable space $(\mathscr{T}, \mathscr{B})$ is called a statistic. The probability measures on $\mathscr{B}$ induced by the statistic T are denoted by $\{P_\theta^T\}, \theta\epsilon\Omega$.

For every $\theta\epsilon\Omega$ and $A\epsilon\mathscr{A}$ there exists an essentially unique real valued $\mathscr{B}$-measurable function $f_\theta(A\,|\,t)$ on $\mathscr{T}$ such that the equation

$$P_\theta(A \cap T^{-1}B) = \int_B f_\theta(A\,|\,t)\, dP_\theta^T \qquad \ldots \quad (1)$$

holds for every $B\epsilon\mathscr{B}$. The set of points t for which $f_\theta(A\,|\,t)$ falls outside the closed interval (0, 1) is of P_θ^T-measure zero for every $\theta\epsilon\Omega$. We call $f_\theta(A\,|\,t)$ the conditional probability of A given that $T = t$ and that θ is the true parameter point.

Definition 2: A statistic T is said to be independent of the parameter θ if, for every $B\epsilon\mathscr{B}$, $P_\theta^T(B)$ is the same for all $\theta\epsilon\Omega$.

Definition 3: The two statistics T and T_1, with associated measurable spaces $(\mathscr{T}, \mathscr{B})$ and $(\mathscr{T}_1, \mathscr{B}_1)$ respectively, are said to be stochastically independent of each other if, for every $B\epsilon\mathscr{B}$ and $B_1\epsilon\mathscr{B}_1$

$$P_\theta(T^{-1}B \cap T_1^{-1}B_1) = P_\theta(T^{-1}B)P_\theta(T_1^{-1}B_1)$$

for all $\theta\epsilon\Omega$.

Now,

$$P_\theta(T^{-1}B \cap T_1^{-1}B_1) = \int_B f_\theta(T_1^{-1}B_1 | t)\, dP_\theta^T.$$

It follows, therefore, that a necessary and sufficient condition in order that T and T_1 are stochastically independent is that the integrand above is essentially independent of t, i.e.

$$f_\theta(T_1^{-1}B_1 | t) = P_\theta(T_1^{-1}B_1) = P_\theta^{T_1}(B_1)$$

for all $t \epsilon \mathcal{T}$ excepting possibly for a set of P_θ^T-measure zero.

Definition 4: The statistic T is called a sufficient statistic (Halmos and Savage, 1949) if for every $A \epsilon \mathcal{A}$ there exists a function $f(A | t)$ which is independent of θ and which satisfies equation (1) for every $\theta \epsilon \Omega$.

Let G be the class of all real valued, essentially bounded, and $\mathcal{B}$-measurable functions on $\mathcal{T}$.

Definition 5: The family of probability measures $\{P_\theta^T\}$ is said to be boundedly complete (Lehmann and Scheffé, 1950) if for any $g \epsilon G$ the identity

$$\int_{\mathcal{T}} g(t) dP_\theta^T \equiv 0 \quad \text{for all } \theta \epsilon \Omega \qquad \ldots \quad (2)$$

implies that $g(t) = 0$ excepting possibly for a set of P_θ^T-measure zero for all θ. $\{P_\theta^T\}$ is called complete if the condition of essential boundedness is not imposed on the integrand in (2). The statistic T is called complete (boundedly complete) if the corresponding family of measures $\{P_\theta^T\}$ is so.

3. Sufficiency and independence

For any two statistics T_1 and T we have for any $B_1 \epsilon \mathcal{B}_1$

$$P_\theta^{T_1}(B_1) = P_\theta(T_1^{-1}B_1) = \int_{\mathcal{T}} f_\theta(T_1^{-1}B_1 | t) dP_\theta^T. \qquad \ldots \quad (3)$$

Now if T be a sufficient statistic then the integrand is independent of θ and if, moreover, T_1 is stochastically independent of T then the integrand is essentially independent of t also. Thus, the right hand side of (3) is independent of θ and so we have

Theorem 1: *Any statistic T_1 stochastically independent of a sufficient statistic T is independent of the parameter θ.*

That the direct converse of the above result is not true will be immediately apparent if we take for the sufficient statistic T the identity mapping of $(\mathfrak{S}, \mathcal{A})$ into itself. No statistic T_1 independent of θ will then be stochastically independent of T excepting in the trival situation where T_1 is essentially equal to a constant. We, however, have the following weaker but important converse.

Theorem 2: *If T be a boundedly complete sufficient statistic then any statistic T_1 which is independent of θ is stochastically independent of T.*

Proof: Since T is sufficient the integrand in (3) is independent of θ. It is also essentially bounded. Now the left hand side of (3) is independent of θ since T_1 is independent of θ. Hence, from boundedly completeness of $\{P_\theta^T\}$ it follows that the integrand in (3) is essentially independent of t as well. That is, T_1 is stochastically independent of T.

In the next section we demonstrate how the above theorem may be used to get a few interesting results in distribution theory.

4. Some characterizations of distributions with location and scale parameters

Let $\underline{x} = (x_1, x_2, ..., x_n)$ be a random variable in an n-dimensional Euclidean space whose probability distribution involves an unknown location parameter μ and a scale parameter $\sigma > 0$. Then any measurable function $f(x_1, x_2, ..., x_n)$ which is independent of change of origin and scale, i.e.

$$f\left(\frac{x_1-a}{b}, ..., \frac{x_n-a}{b}\right) \equiv f(x_1, ..., x_n)$$

for all a and $b > 0$ is independent of the unknown parameter (μ, σ). Now, if there exists a boundedly complete sufficient statistic T for (μ, σ) then f must be stochastically independent of T. For example, if $x_1, x_2, ..., x_n$, are independent observations on a normal variable with mean μ and s.d. σ then it is well known that $T = (\bar{x}, s)$ is a sufficient statistic ($\bar{x}$ is the sample mean and s the sample s.d.). The completeness of T follows from the unicity property of the bivariate Laplace transform. It then follows from Theorem 2 that any measurable function $g(\bar{x}, s)$ of $\bar{x}$ and s is stochastically independent of any measurable function $f(x_1, x_2, ..., x_n)$ of the observations that is independent of change of origin and scale. The functions g and f need not be real valued. For instance, we may have

$$g = (\Sigma x_i^2, \sum_{(i \neq j)} x_i x_j)$$

and

$$f = \left(\frac{\Sigma(x_i - \bar{x})^3}{s^3}, \frac{\Sigma(x_i - \bar{x})^4}{s^4}, ...\right).$$

Again the stochastic independence of $\bar{x}$ and s follows from the fact that, for any fixed σ, the statistic $\bar{x}$ is a complete sufficient statistic for μ and that s, by virtue of its being independent of change of origin, is independent of the location parameter μ.

Now let $x_1, x_2, ..., x_n$, be independent random variables with joint d.f. $F_1(x_1 - \theta), F_2(x_2 - \theta), ..., F_n(x_n - \theta)$.* Since θ is a location parameter it follows that any linear function $\Sigma a_i x_i$ with $\Sigma a_i = 0$ is independent of θ. If $\Sigma b_i x_i$ is a boundedly complete sufficient statistic for θ then from Theorem 2 it follows that $\Sigma a_i x_i$ is independent of $\Sigma b_i x_i$.

* For the sake of notational convenience, we make no distinction between random variables and the values that they may assume.

Now, since $\Sigma b_i x_i$ is a sufficient statistic it follows that every $b_i \neq 0$. For, if possible, let $b_j = 0$. Then x_j is stochastically independent of $\Sigma b_i x_i$ and so from Theorem 1 x_j is independent of the parameter θ which contradicts the assumption that the d.f. of x_j is $F_j(x_j - \theta)$. Again, we can take all the a_i's different from zeros. Thus, the two linear functions $\Sigma a_i x_i$ and $\Sigma b_i x_i$ (with non-zero coefficients) of the independent random variables $x_1, x_2, \ldots, x_n$, are stochastically independent. Therefore,† all the x_i's must be normal variables. We thus have the following:

Theorem 3: *If $x_1, x_2, \ldots, x_n$, are independent random variables such that their joint d.f. involves an unknown location parameter θ then a necessary and sufficient condition in order that $\Sigma b_i x_i$ is a boundedly complete sufficient statistic for θ is that $b_i > 0$ and that x_i is a normal variable with mean θ and variance b_i^{-1} $(i = 1, 2, \ldots, n)$.*

Let us now turn to the case of the Gamma variables. Let $x_1, x_2, \ldots, x_n$, be independent Gamma variables with the same scale parameter $\theta > 0$, i.e., the density function of x_i is

$$f_i(x)dx = \frac{1}{\Gamma(m_i)\theta^{m_i}} x^{m_i - 1} e^{-x/\theta} dx \qquad (x \geqslant 0, \quad \theta > 0, \quad m_i > 0).$$

It is clear then that Σx_i is a sufficient statistic for θ and its completeness follows from the unicity property of the Laplace transform. Thus, we at once have the well known result that Σx_i is stochastically independent of any function $f(x_1, x_2, \ldots, x_n)$ that is independent of change of scale (i.e. independent of θ).

Recently it has been proved by R. G. Laha that if $x_1, \ldots, x_n$, are independent and identically distributed chance variables and if Σx_i is independent of $\Sigma a_{ij} x_i x_j / (\Sigma x_i)^2$ then (under some further assumptions) all the x_i's must be Gamma variables. Using this result we can immediately get a characterization of the Gamma distribution analogous to Theorem 3.

REFERENCES

BASU, D. (1951): On the independence of linear functions of independent chance variables. *Bull. Int. Stat. Inst.*, 33, Pt. 2, 83-96.

DARMOIS, G. (1951): Sur diverses propriétés caractéristiques de la loi de probabilité de Laplace-Gauss, *Bull. Int. Stat. Inst.*, 33, Pt. 2, 79-82.

——— (1953): Analyse générale des liaisons stochastiques—Etude particulière de l'analyse factorielle linéaire. *Rev. Inst. Internat. Stochastique*, 21, 2-8.

HALMOS, P. R. AND SAVAGE, L. J. (1949): Application of the Radon-Nikodym theorem to the theory of sufficient statistics. *Ann. Math. Stat.*, 20, 225-241.

LEHMANN, E. L. AND SCHEFFÉ, H. (1950): Completeness, similar regions and unbiased estimation. *Sankhyā*, 10, 305-340.

LAHA, R. G. (1954): On a characterization of the Gamma distribution. *Ann. Math. Stat.*, 25, 784-787.

† This result was first conjectured (and proved under certain assumptions) by the author in 1951. The proof without any assumption is due to G. Darmois (1953).