DENSITY UNBIASED POINT ESTIMATES

By Raymond P. Peterson

University of Washington and University of California, Riverside

1. Summary. A new concept of unbiasedness (density unbiasedness) for point estimates is introduced and the "best" density unbiased point estimate for the mean of any normal distribution is proved to be the ordinary sample mean $\bar{x} = \sum_{i=1}^{n} x_i/n$. Under certain conditions on the form of the characteristic function of a family of probability density functions involving an unknown location parameter, $\bar{x}$ is shown to be a density unbiased point estimate of the location parameter.

2. Introduction. Let $x_1, \cdots, x_n$ denote n (not necessarily independent) random variables each of which is distributed over a space M according to a distribution function $P(x, \theta, \theta_1, \cdots, \theta_s)$. It is assumed that $P(x, \theta, \theta_1, \cdots, \theta_s)$ is completely specified except for the $s + 1$ parameters $\theta, \theta_1, \cdots, \theta_s$. These parameters may be represented by a point (θ, σ) in the $(s + 1)$-dimensional parameter space Ω_{s+1} where $\sigma = (\theta_1, \cdots, \theta_s)$. Also $X_n = (x_1, \cdots, x_n)$ is a point in the n-dimensional sample space M_n. We shall assume here that $P(x, \theta, \sigma)$ is absolutely continuous and denote the corresponding probability density function by $p(x, \theta, \sigma)$. Let $p_n(X_n, \theta, \sigma) = p_n(x_1, \cdots, x_n, \theta, \sigma)$ denote the joint probability density function at the point $X_n \varepsilon M_n$.

A statistical point estimate of the parameter θ, which ranges over a subset ω of the real line, is a function $f(X_n)$ of the sample values $x_1, \cdots, x_n$, whose range is the same subset ω.

Let $\mathfrak{F}$ be the $(s + 1)$-parameter family of probability density functions

$$p(x, \theta, \sigma) = p(x, \theta, \theta_1, \cdots, \theta_s).$$

The *mean probability density function of* $\mathfrak{F}$ *generated by* $f(X_n)$ *relative to* θ is given by

$$\varphi_f(x, \theta, \sigma) = \int_{M_n} p(x, f(X_n), \sigma)\, p_n(X_n, \theta, \sigma)\, dX_n. \tag{2.1}$$

It is readily seen that $\varphi_f(x, \theta, \sigma)$ is a probability density function provided $p(x, f(X_n), \sigma)$ is measurable in X_n over M_n.

A point estimate $f(X_n)$ of θ will be called *density unbiased* if

$$\varphi_f(x, \theta, \sigma) \equiv p(x, \theta, \sigma'),$$

where σ' is some value of σ. There are various criteria by which we might choose a "best" density unbiased estimate from the class of all density unbiased estimates, provided this class is not empty. We shall call an estimate $f(X_n)$ of θ a *best density unbiased estimate* if

i) $f(X_n)$ is a density unbiased point estimate of θ, and

Received 1/13/53, revised 11/15/53.

ii) $(\sigma' - \sigma)^2 = \sum_{i=1}^{s} (\theta_i - \theta_i)^2$ is minimized by $f(X_n)$ with respect to all density unbiased estimates of θ.

3. The best density unbiased point estimate for the mean of a normal population. Let $\mathfrak{F}$ denote the two-parameter family of probability density functions

$$p(x, \theta, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\theta/\sigma)^2/2}$$

and let $f(X_n)$ be any point estimate of θ where $x_1, \cdots, x_n$ are n independent sample values of x. Now $f(X_n)$ will be density unbiased if and only if $\varphi_f(x, \theta, \sigma)$ possesses a characteristic function of the form

$$e^{(t\sigma')^2/2}. \tag{3.1}$$

THEOREM 3.1. *An estimate $f(X_n)$ is a density unbiased estimate of the mean θ of a normal distribution if and only if $f(X_n)$ is itself normally distributed with mean θ.*

PROOF. Taking the characteristic function of both sides of the identity (2.1) we have

$$\begin{aligned}\psi_f(t, \theta, \sigma) &= \int_{-\infty}^{\infty} e^{itx}\varphi_f(x, \theta, \sigma)\, dx \\ &= \int_{-\infty}^{\infty} e^{itx} \int_{M_n} p(x, f(X_n), \sigma)\, p_n(X_n, \theta, \sigma)\, dX_n\, dx \\ &= \int_{M_n} e^{itf(X_n)-t^2\sigma^2/2} p_n(X_n, \theta, \sigma)\, dX_n.\end{aligned}$$

However, $\psi_f(t, \theta, \sigma)$ must be of the form (3.1). Therefore

$$\int_{M_n} e^{itf(X_n)} p_n(X_n, \theta, \sigma)\, dX_n \equiv e^{it\theta - t^2(\sigma'^2-\sigma^2)/2}. \tag{3.2}$$

On the other hand

$$\int_{M_n} e^{itf(X_n)} p_n(X_n, \theta, \sigma)\, dX_n \equiv \int_{-\infty}^{\infty} e^{itf} g(f, \theta, \sigma)\, df \tag{3.3}$$

where $g(f, \theta, \sigma)$ is the probability density function of $f(X_n)$. Since the right-hand side of equation (3.3) is the characteristic function of $f(X_n)$, it follows from (3.2), (3.3) and the uniqueness theorem for characteristic functions [1] that $f(X_n)$ must be normally distributed with mean θ and variance $\sigma'^2 - \sigma^2$.

It follows from Theorem 3.1 that in looking for a best density unbiased estimate of θ we can restrict ourselves to the class N_θ of normally distributed estimates with mean θ. Now, the ordinary sample mean $\bar{x}$ is known to be normally distributed with mean θ and minimum variance σ^2/n among all estimates with mean θ, subject to certain regularity conditions which are satisfied by all estimates in N_θ. Therefore $\sigma'^2 - \sigma^2$ is a minimum for $f(X_n) = \bar{x}$.

Since $\sigma' > \sigma$, and $\bar{x}$ minimizes $\sigma'^2 - \sigma^2$, it immediately follows that $\bar{x}$ minimizes $(\sigma' - \sigma)^2$ with respect to the class of all density unbiased estimates of θ. Hence $\bar{x}$ is the best density unbiased estimate of the mean θ.

4. Density unbiased point estimates of a location parameter. Let $\mathfrak{F}$ denote the two-parameter family of probability density functions

$$p(x, \theta, \sigma) = \frac{1}{\sigma} p\left(\frac{x - \theta}{\sigma}\right), \qquad \sigma > 0$$

and let $x_1, \cdots, x_n$ denote n independent observed values of x. In this case θ is called a *location* parameter and σ a *scale* parameter.

LEMMA 4.1. *The characteristic function $\psi(t, \theta, \sigma)$ of $p(x, \theta, \sigma)$ is of the form $\psi(t, \theta, \sigma) = e^{it\theta}h(t\sigma)$, where $h(t\sigma) = \psi(t\sigma, 0, 1)$.*

PROOF.

$$\psi(t, \theta, \sigma) = \frac{1}{\sigma} \int_{-\infty}^{\infty} e^{itx} p\left(\frac{x - \theta}{\sigma}\right) dx = e^{it\theta} \int_{-\infty}^{\infty} e^{it\sigma z} p(z)\, dz$$

where $z = (x - \theta)/\sigma$. If we let $h(t\sigma) = \int_{-\infty}^{\infty} e^{it\sigma z} p(z)\, dz$, the lemma follows.

THEOREM 4.1. *If $h(t\sigma_1)h(t\sigma_2) \equiv h(t\sigma_3)$, where σ_1, σ_2, and σ_3 are values of σ independent of t, then $\bar{x}$ is distributed according to a member of $\mathfrak{F}$.*

PROOF.

$$\begin{aligned} E(e^{it\bar{x}}) &= \int_{-\infty}^{\infty} e^{it\bar{x}}\, p_n(X_n, \theta, \sigma)\, dX_n = \prod_{j=1}^{n} \int_{-\infty}^{\infty} e^{itx_j/n} p(x_j, \theta, \sigma)\, dx_j \\ &= e^{it\theta}[h(t\sigma/n)]^n = e^{it\theta}h(t\sigma'). \end{aligned} \tag{4.1}$$

Hence by Lemma 4.1 the probability density function of $\bar{x}$ belongs to $\mathfrak{F}$.

THEOREM 4.2. *If $h(t\sigma_1)h(t\sigma_2) \equiv h(t\sigma_3)$, where σ_1, σ_2, and σ_3 are values of σ independent of t, then $\bar{x}$ is a density unbiased estimate of the location parameter θ.*

PROOF.

$$\begin{aligned} \psi_{\bar{x}}(t, \theta, \sigma) &= \int_{-\infty}^{\infty} e^{itx} \varphi_{\bar{x}}(x, \theta, \sigma)\, dx = \int_{-\infty}^{\infty} e^{itx} \int_{M_n} p(x, \bar{x}, \sigma) p_n(X_n, \theta, \sigma)\, dX_n\, dx \\ &= \int_{M_n} e^{it\bar{x}} h(t\sigma) p_n(X_n, \theta, \sigma)\, dX_n = h(t\sigma) E(e^{it\bar{x}}). \end{aligned}$$

From (4.1) we have $E(e^{it\bar{x}}) = e^{it\theta}h(t\sigma')$ and so

$$\psi_{\bar{x}}(t, \theta, \sigma) = e^{it\theta}h(t\sigma)\, h(t\sigma') = e^{it\theta}h(t\sigma'').$$

Therefore $\varphi_{\bar{x}}(x, \theta, \sigma) \,\varepsilon\, \mathfrak{F}$ and the theorem follows directly.

EXAMPLE. Let $\mathfrak{F}$ denote the two-parameter family of Cauchy distributions given by

$$p(x, \theta, \sigma) = \frac{1}{\pi\sigma} \frac{1}{1 + (x - \theta)^2/\sigma^2}, \qquad \sigma > 0.$$

In this case the characteristic function $\psi(t, \theta, \sigma) = e^{i\theta t - \sigma|t|}$, so that $h(t\sigma) = e^{-\sigma|t|} = e^{-|\sigma t|}$. It is immediately verified that $h(t\sigma_1)\, h(t\sigma_2) = h(t\sigma_3)$, where $\sigma_3 = \frac{1}{2}(\sigma_1 + \sigma_2)$. Therefore, by Theorem 4.1, $\bar{x}$ has a Cauchy distribution. In fact, since in this case $E(e^{it\bar{x}}) = [h(t\sigma/n)]^n = h(t\sigma)$, $\bar{x}$ has exactly the same distribution as does x itself, namely $p(x, \theta, \sigma)$. It follows from Theorem 4.2 that $\bar{x}$ is a density unbiased point estimate of the location parameter θ.

It is readily seen that Theorems 4.1 and 4.2 are valid also when $\bar{x}$ is replaced by any linear homogeneous estimate $\sum_{i=1}^{n} a_i x_i$, where $\sum_{i=1}^{n} a_i = 1$.

REFERENCE

[1] HARALD CRAMÉR, *Mathematical Methods of Statistics*, Princeton University Press, 1946.

SOME PROPERTIES OF BETA AND GAMMA DISTRIBUTIONS

BY M. V. JAMBUNATHAN

University of Mysore

1. Summary and Introduction. The object of this paper is to present certain important properties of the Gamma distribution and the two kinds of Beta distributions, and to indicate certain useful applications of these two to sampling problems. The distribution of the Studentised D^2-statistic under the null hypothesis is obtained in two different ways.

2. The Gamma distribution. If a random variable x has probability density

$$\frac{1}{\Gamma(a)} e^{-x} x^{a-1}, \qquad 0 \leqq x < \infty, \tag{2.1}$$

then x is said to have a *Gamma distribution*; furthermore, x is called a *Gamma variate* with parameter a, and is symbolically written $\gamma(a)$. The Gamma distribution is known to possess, among others, the mean conserving property (m.c.p.), provided the variates are independent.[1] Symbolically

$$\gamma(a) + \gamma(b) = \gamma(a + b).$$

If x is a Gamma variate with parameter a, then $2x$ is distributed as χ^2 with $2a$ degrees of freedom.

3. Beta distribution of the first kind. A random variable having the probability density

$$\frac{1}{B(a, b)} x^{a-1}(1 - x)^{b-1}, \qquad 0 \leqq x \leqq 1, \tag{3.1}$$

Received 9/25/52, revised 1/20/54.

[1] This necessary proviso which was left out in the original paper has since been inserted at the suggestion of a referee to whom the author is thankful for several other suggestions besides.