1. ARTÍCULOS DE ESTADÍSTICA

WHY (NOT) FREQUENTIST INFERENCE (TOO)?

George Casella [1]
University of Florida

1. Introduction

It was very kind of Prof. Gómez Villegas to invite me to write an article with the title "Why Frequentist Inference". However, when thinking things over, it occurred to me that there is no reason to take the position of defending frequentist inference, or in trying to convince anyone that frequentist inference is the only correct way to do statistics. To take such a position is like trying to convince Raúl González, a star of the Spanish World Cup Soccer Team, to kick goals with only his right foot.

The position that I am comfortable with is to view frequentist inference as one part of the tool kit of the complete statistician. Some things are best done using Bayesian tools, some things are best done using frequentist tools, and the complete statistician uses all of the available tools. (We also do not forget that there are likelihood tools, nonparametric tools, etc.)

2. Construction or Evaluation

The frequentist approach, by its very being, is an evaluative methodology. That is, the frequentist concept of repeated trials, where we measure performance by long run averages, gives us a setting in which we can evaluate any procedure that we are considering using. However, in no part of frequentist statistics does it tell us how to *construct* a procedure.

Other than obvious cases where sample means and variances are good candidates for estimates of parameters of interest, there is no "recipe" or set of rules that a frequentist can use to construct an estimator. For example, if we observe a sample $y_1, y_2, \ldots, y_n$ from a population with mean μ and variance σ^2, we might consider

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

"frequentist" estimators. But there is nothing frequentist about the construction! The long-run-evaluative framework did not *give* us the estimators, it only allows us to evaluate them. The fact that s^2, where we divide by $n-1$, is unbiased, is a response to a frequentist evaluation - the property of unbiasedness - but it is not a construction. There is no frequentist rule for constructing an unbiased estimator. (The closest we can come is the Rao-Blackwellization argument; see Casella and Berger 2001, Section 7.3.3), but that does not tell us how to construct an unbiased estimator. It tells us how to make a given unbiased estimator better.

As an example of this point, suppose $y_1, y_2, \ldots, y_n$ is a sample from a gamma density $F(y|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1}e^{-y/\beta}$, and we are interested in an estimator of α. Question: What frequentist construction will get us an estimator of α? Answer: None.

Of course, we can use method of moments, maximum likelihood, Bayes, or even good guessing to get an estimator. And once we have an estimator we can *then* use frequency theory to evaluate its worth. For a point estimator we can evaluate its bias, its mean squared error, or any other type of risk function that we desire. We can ask if the estimator is admissible, minimax, unbiased, equivariant, consistent, efficient, or has any of a multitude of other desirable properties. If we construct an interval estimate we can assess its coverage probability, optimize its length, or compare it to other interval estimators.

3. So We Do Frequentist Inference "Too"

In his ASA President's Invited Address, Rod Little (Little 2006) argued for a "Calibrated Bayes" approach, in which we use Bayesian methods for inference in a particular model, but frequentist methods for model assessment. He cites other prominent Bayesians (Jim Berger, Don Rubin) who ha-

[1]Distinguished Professor, Department of Statistics, University of Florida, Gainesville, FL 32611-8545. Supported by National Science Foundation Grant DMS-04-05543. Email: casella@stat.ufl.edu

ve similar opinions. To me, this is a fine way to do statistics, and I have been a proponent of such an approach for a long time. The rise of MCMC has made such an arrangement the standard way of doing business. In its construction MCMC, especially if implemented through a Gibbs sampler, is a Bayesian solution to a problem. To assess the solution we monitor the MCMC output which, of course, is the monitoring of many repeated trials, and hence is frequentist. This is the perfect example of how the two approaches complement each other. The extremely flexible Bayesian models can help us get solutions to many complex problems, and the frequentist evaluation tools help us evaluate and calibrate our inferences.

Thus, we should be neither Bayesians nor frequentists. We should, perhaps, be "calibrated Bayesians" according to Little, but we really should be more. When faced with a problem, we should use all the tools available (Bayes, frequentist, likelihood), and use each in their best possible way, to arrive at the best solution that we can give for a particular problem. Then, when we truly use all of our tools we are neither frequentist nor Bayesians, we are statisticians!

Referencias

[1] Casella, G. and Berger, R. L. (2001). *Statistical Inference, Second Edition.* Monterey: Duxbury.

[2] Little, R. J. (2006). Calibrated Bayes: A Bayes/Frequentist Roadmap. *American Statistician*, **60**, 213-223.

CONJUNTOS ALEATORIOS: ESPERANZAS DE AUMANN Y HERER

Pedro Terán
Grupo Decisión Multicriterio Zaragoza
Unidad Docente de Métodos Estadísticos
Universidad de Zaragoza

Resumen

En este artículo presentamos una síntesis del trabajo *Relaciones entre las esperanzas de Aumann y Herer de un conjunto aleatorio* (Premio Ramiro Melendreras 2006).

1. Introducción

Por así decirlo, los conjuntos aleatorios "siempre han estado ahí"; ya Kolmogorov, en su libro frecuentemente citado como el origen de la teoría moderna de la probabilidad [12], discute la idea de *una región cuya forma depende del azar*.

La aguja que se lanza al azar en el problema de Buffon es un ejemplo de conjunto aleatorio. Un proceso de Poisson, o en general cualquier proceso de puntos como los ceros de un movimiento browniano, los instantes en que un proceso continuo rebasa una barrera o un proceso de records, es un conjunto aleatorio. Los puntos óptimos en un problema de optimización estocástica no convexa (es decir, con solución en general no única) forman un conjunto aleatorio. Un intervalo o una región de confianza, así como el resultado de un algoritmo de análisis *cluster*, son conjuntos aleatorios.

Pero, pese a esta abundancia de ejemplos estadísticos, el desarrollo de la teoría de conjuntos aleatorios ha venido motivado sobre todo por aplicaciones en áreas como la estereología (en Geología), la teoría económica, la geometría estocástica, el análisis de imágenes y otras.

Uno de los conceptos fundamentales que se han afrontado desde el punto de vista particular de varias de esas áreas es el de **esperanza** de un conjunto aleatorio. En el caso de variables aleatorias, la esperanza se calcula usando la integral de Lebesgue. Es un leve defecto, que solemos considerar anecdótico, el que el valor esperado de la tirada de un dado sea 3,5 –un valor que de hecho nadie espera. Pero este mismo fenómeno constituye una dificultad enorme para encontrar una buena definición de esperanza de un conjunto aleatorio: está claro que la media de dos círculos debe ser un círculo de posición y tamaño intermedios, pero ¿está claro cuál "debería" ser la media de un círculo y un triángulo, de un pen-