NIH Public Access
Author Manuscript
Commun Stat Theory Methods. Author manuscript; available in PMC 2013 December 16.

Published in final edited form as:
Commun Stat Theory Methods. 2001 January 1; 30(2): . doi:10.1081/STA-100002037.

On the Expected Values of Sequences of Functions

Deborah H. Glueck[1] and **Keith E. Muller**[2]

[1]Dept. of Prev. Med. and Biometrics, University of Colorado HSC, Campus Box B119, 4200 East Ninth Avenue, Denver, CO 80262

[2]Dept. of Biostatistics, University of North Carolina, Chapel Hill, NC 27599-7400

Abstract

We prove new extensions to lemmas about combinations of convergent sequences of distribution functions and absolutely continuous bounded functions. New lemma one, a generalized Helly theorem, allows computing the limit of the expected value of a sequence of functions with respect to a sequence of measures. Previously published results allow either the function or the measure to be a sequence, but not both. Lemma two allows computing the expected value of an absolutely continuous monotone function by integrating the probabilities of the inverse function values. Previous results were restricted to the identity function. Lemma three gives a computationally and analytically convenient form for the limit of the expected value of a sequence of functions of a sequence of random variables. This is a new result that follows directly from the first two lemmas. Although the lemmas resemble standard results and seem obviously true, we have found only similar looking and related but quite distinct results in the literature. We provide examples which highlight the value of the new results.

Keywords

Absolutely continuous; Inversion; Integrals

1. Introduction

Computing expectations, both analytically and numerically, remains one of the central problems in statistics. We present three new lemmas which aid both analytic and numerical calculation for some applications in small and large samples.

In particular, the first lemma allows calculation of the limits of expectations, when both the function and the random variable are converging to limits. The second lemma suggests natural transformations for the computation of expectations, a common statistical task. We suggest transformations based on probability distribution functions and their inverses. The required numerical functions are widely available in common statistical programming languages. The choice automatically simplifies numerical computation of expectations by leading to evaluating bounded functions on bounded regions. Although a transformation is a standard numerical technique, the use of probability functions eliminates the guess work involved in choosing the transformation, and provides an ideal match between statistical thinking and ease of computation. The third lemma combines the results and allows one to calculate an expectation when both the function and the random variable are converging to a limit.

The need for the results presented here arose from a desire to derive computable expressions for the power of certain hypothesis tests in multivariate regression with both fixed and random predictors. We provide examples of how we used the lemmas to numerically calculate small sample expectations, and to correctly find limiting results.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Calculating the expected value of variables and functions can be difficult, especially in the limit. Often either the function of integration, or the limits, or both, are unbounded. In attempting to derive large sample properties, a sequence of cumulative distribution functions (CDF's) may converge to a point mass, while their support converges to a set of measure zero. The Riemann integral fails under these conditions. Thus the lemmas that follow must be stated in terms of Lebesgue integrals computed with respect to probability measures.

Although the lemmas resemble standard results and seem obviously true, we have found only similar looking and related but quite distinct results in the literature. All depend on various restrictive regularity conditions that often arise in practice. For example, for a sequence of cumulative distribution functions, $\{F_n\}$, Serfling (1, p16) proved that if $F_n \Rightarrow F$ then for any bounded continuous function g,

$$\lim_{n\to\infty}\int g\ dF_n=\int g\ dF. \quad (1)$$

Pratt (2, p74) and Loeve (3, pl26) gave results similar to Equation 1. We give a more general result in which both the integrand and the measure are converging. Gibbons and Chakraborti (4, p37-38) mentioned a quantile transform result, without proof. They define the quantile function using the *infimum*. Our Corollary 3 is a special case of their result. We illustrate the value of the new results with three examples concerning power of certain multivariate tests.

2. Three Lemmas

Lemma 1

Consider a random variable, X, and a sequence of random variables, $\{X_1, X_2, \ldots\}$, with corresponding CDFs F, and $\{F_1, F_2, \ldots\}$. Suppose F_n converges to F, and thus X_n converges in distribution to X. Let $\{g_1(x), g_2(x), \ldots\}$ be a set of continuous bounded functions such that $g_n(x)$ converge uniformly to g(x), a continuous bounded function. Assume that $\forall n \int g_n\, dF_n < \infty$ and $\int g\, dF < \infty$. Taking the integrals with respect to the Lebesgue-Stieltjes probability measures induced by F and $\{F_1, F_2, \ldots\}$ (5, p69),

$$\lim_{n\to\infty}\int g_n dF_n=\int g\ dF. \quad (2)$$

The special case of $\int g d F_n$ corresponds to a Helly theorem, (5, p192–194). Also, the results of exercise 9-2 in Burrill (5, p195) indicate that less stringent regularity conditions would be hard to find for the special case of $\int g_n d F$.

Proof. It suffices to show that $\forall \in > 0$, $\exists\, M(\in) > 0$ such that for $n > M(\in)$,

$$\left|\int g_n d\ F_n-\int g\ dF\right|<\epsilon. \quad (3)$$

By assumption, $g_n(x)$ converges uniformly to $g(x)$. Thus, $\forall \in > 0$, there exists a corresponding number $M_1(\in) > 0$ such that for all $n > M_1$ and for all x in the domain of g, $|g_n(x) - g(x)| < \in/2$ (6, p530). Thus, for $n > M_1(\in)$,

$$\left|\int g_n d\ F_n-\int g\ dF_n\right|=\left|\int (g_n-g) d\ F_n\right|\le \left|\int |g_n-g| d\ F_n\right|<\left|\int |\epsilon/2| d\ F_n\right|\le |\epsilon/2|\left|\int d\ F_n\right|\le \epsilon/2. \quad (4)$$

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The last step follows because F_n is a cumulative probability distribution function, and hence $\forall_n$, $\int d\, F_n = 1$. By part 3 of a theorem in Serfling (1, p16), $\forall \in > 0$, we may conclude that $\exists$ $M_2(\in) > 0$ such that for $n > M_2(\in)$,

$$|\int g\ dF_n - \int g\ dF| < \epsilon/2. \quad (5)$$

Now, $\forall \in > 0$ and $\forall x$ in the domain of g, choose $n > \max[M_1(\in), M_2(\in)]$. Then

$$|\int g_n dF_n - \int g\ dF| < \epsilon/2 + \epsilon/2 = \epsilon, \quad (6)$$

with the inequality following from the triangle inequality.

Lemma 2

(Corollary to Lemma 2.1, 7, p243) Let X be a continuous random variable with density $f_x(x)$ and distribution function $F_X(x)$. Let g(x) be a real valued absolutely continuous function that is strictly monotone decreasing in x, so that g(x) > y iff x < $g^{-1}(y)$. Let

$$\begin{aligned} A &= \{y: y \geq 0\} \cap \{y: P\{g(X) > y\} > 0\}\ and \\ \mathscr{B} &= \{y: (y > 0)\} \cap \{y: P\{g(X) < -y\} > 0\}. Then \\ \varepsilon[g(X)] &= \int_A F_X[g^{-1}(y)]dy - \int_{\mathscr{B}} \{1 - F_X[g^{-1}(-y)]\}dy. \end{aligned} \quad (7)$$

Proof. Note that

$$\begin{aligned} \varepsilon[g(X)] &= \int_A P\{g(X) > y\}dy \\ &\quad - \int_{\mathscr{B}} P\{g(X) < \\ &\quad - y\}dy \\ &= \int_A P\{X < g^{-1}(y)\}dy - \int_{\mathscr{B}} P\{X > g^{-1}(\\ &\quad - y)\}dy \\ &= \int_A P\{X < g^{-1}(y)\}dy \\ &\quad - \int_{\mathscr{B}} (1 \\ &\quad - P\{X < g^{-1}(-y)\})dy. \end{aligned} \quad (8)$$

The result follows.

Corollary 2.1—With the same conditions as in Lemma 1, and for b > a > 0, suppose $\forall x \in \mathscr{R}$, $g(x) \in [a, b]$. Then

$$\varepsilon[g(X)] = \int_a^b F_X[g^{-1}(y)]dy. \quad (9)$$

Corollary 2.2—With the same conditions as in Lemma 1, consider instead h(x), a real valued absolutely continuous function that is strictly monotone increasing in x, so that h(x) > y iff x > $h^{-1}(y)$. Then

NIH-PA Author Manuscript

$$\varepsilon[h(X)]=\int_{A}\left\{1-F_X[h^{-1}(y)]\right\}dy-\int_{\mathscr{B}}F_X[h^{-1}(-y)]dy. \quad (10)$$

Lemma 3

Consider the continuous random variable, X, and the sequence of continuous random variables, $\{X_1, X_2, \ldots\}$, with the same assumptions as Lemma 1.Let g(x)and the set $\{g_1(x), g_2(x), \ldots\}$ be real valued absolutely continuous bounded functions that are strictly monotone decreasing in x, so that $g_n(x) > y$ iff $x<g_n^{-1}(y)$. Suppose the sequence $\{g_1(x), g_2(x), \ldots\}$ converges uniformly to g(x). Assume that $\int g\, dF < \infty$ and $\forall n, \int g_n\, dF_n < \infty$. For $b > a > 0$, suppose $\forall x \in \mathscr{R}$, $g(x) \in [a, b]$. Then

$$\lim_{n\to\infty}\varepsilon[g_n(X_n)]=\int_a^b F_X[g^{-1}(y)]dy. \quad (11)$$

Proof. Follows directly from Lemma 1 and Lemma 2.

Corollary 3—(Quantile transformation: see 4, §2.5, p37–38) Consider a real valued random variable X with strictly monotone distribution function $F_x(x)$ and density function $f_x(x)$, defined on the interval (a, b), with $-\infty < a < b < \infty$. Let $y = F_x(x)$. Then

$$\varepsilon(X)=\int_0^1 F_X^{-1}(y)dy. \quad (12)$$

When resorting to numerical techniques for calculating expectations, either the density or the region of integration, or both, may be infinite. This transformation reduces the problem to an integral of a bounded function over a bounded interval.

3. Examples

Example for Lemma 1

Glueck (8) considered taking the limit, under a sequence of Pitman local alternatives, of an approximation for the power of the Hotelling-Lawley trace statistic, with Gaussian predictors. The asymptotic power can be written as the expected value of a non-central *F*, with respect to the distribution of a random noncentrality value. In this setting, the Riemann integral is undefined because the support for the random noncentrality parameter converges to set of measure zero as the parameter converges to a point. Let $F(\nu_1, \nu_{2N}, \omega_N)$ indicate a noncentral *F* random variable, with denominator degrees of freedom and noncentrality depending on *N*. Suppose

$$g_N=\Pr\left\{F\left(\nu_1,\nu_{2N},\omega_N\right)\leq f_N\right\}. \quad (13)$$

Consider integrating g_N with respect to F_n, the distribution function of a sum of independent scaled χ^2 random variables for which the scaling constants depend on ω_N, and the degrees of freedom depend on *N*. With α the type 1 error rate, c_{crit} chosen so that $F_{\chi^2}(c_{crit}; ab) = 1 - \alpha$, and $\lim_{N\to\infty}$, $\omega = \omega_L$,

$$\lim_{N\to\infty}\int g_N dF_N=1-F_{\chi^2}(c_{\text{crit}};\nu_1,\omega_L). \quad (14)$$

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Example for Corollary 2.1

Glueck (8) sought a computational form for the small sample power of the Hotelling-Lawley trace statistic, with Gaussian predictors. Suppose $F_{\omega}(w)$ is the distribution function of ω, a sum of independent scaled χ^2 random variables. Define

$$g(\omega) = \Pr\{F(\nu_1, \nu_2, \omega) < f\}, \quad (15)$$

with $\omega \in [0, \infty]$ and f chosen so that $g(\omega) \in [0, 1 - \alpha]$. Then

$$\varepsilon[g(\omega)] = \int_0^{\infty} g(\omega) dF_{\omega} = \int_0^{1-\alpha} F_{\omega}[F^{-1}(f; \nu_1, \nu_2, y)] dy. \quad (16)$$

Example for Corollary 3

Muller and Pasour (9) defined

$$g(t) = F_{\chi^2}\left(\frac{\nu_{1s}}{\nu_{2s}} f_R t; \nu_{1s}, \omega_s\right) - F_{\chi^2}\left(\frac{\nu_{1s}}{\nu_{2s}} f_L t; \nu_{1s}, \omega_s\right), \quad (17)$$

and considered integrating

$$F_V(z) = \pi_s^{-1} \int_0^{z_*} g(t) f_{\chi^2}(t; \nu_{2s}) dt. \quad (18)$$

They used a particular quantile transformation to produce a much better behaved numerical integral. If $p = F_{\chi^2}(t; \nu_{2s})$, then $t = F_{\chi^2}^{-1}(p; \nu_{2s})$ and $dp = f_{\chi^2}(t; \nu_{2s})dt$. If $p_0 = F_{\chi^2}(z\nu_{2s}/\sigma^2; \nu_{2s})$ then

$$F_V(z) = \pi_s^{-1} \int_0^{p_0} g_R[F_{\chi^2}^{-1}(p; \nu_{2s})] dp - \pi_s^{-1} \int_0^{p_0} g_L[F_{\chi^2}^{-1}(p; \nu_{2s})] dp. \quad (19)$$

Acknowledgments

The authors gratefully acknowledge the help of Drs. G. G. Koch, R. M. Hamer, L. M. LaVange, D. F. Ransohoff and P. W. Stewart. In addition, an anonymous Associate Editor stimulated us to clarify the motivation of the paper, provide more compelling examples, and improve the readability of the paper. Also, an anonymous Associate Editor brought the work of Gibbons and Chakraborti to our attention. Glueck was supported in part by grant number T32HS000589-04 from the Agency for Health Care Policy and Research to the University of Medicine and Dentistry of New Jersey, and by National Cancer Institute Grant IP30CA46934-01 to the University of Colorado. Muller's work supported in part by NCI program project grant P01 CA47 982-04.

References

1. Serfling, RJ. Approximation Theorems of Mathematical Statistics. John Wiley and Sons; New York: 1980.
2. Pratt JW. On interchanging limits and integrals. Annals of Mathematical Statistics. 1960; 31:74–77.
3. Loève, M. Probability Theory II. 4th. Springer-Verlag; New York: 1977.
4. Gibbons, JD.; Chakraborti, S. Nonparametric Statistical Inference. 3rd. Marcel Dekker; New York: 1992.
5. Burrill, CW. Measure, Integration and Probability. McGraw-Hill; New York: 1972.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

6. Courant, R.; John, F. Introduction to Calculus and Analysis, Volume One. Interscience Publishers; New York: 1965.
7. Ross, S. A First Course in Probability. 2nd. Macmillan; New York: 1984.
8. Glueck, DH. Power for a Generalization of the General Linear Multivariate Model with Fixed and Random Predictors, Mimeo Series No 2158T. Institute of Statistics; Chapel Hill, North Carolina: 1996.
9. Muller KE, Pasour VB. Bias in linear model power and sample size due to estimating variance. Communications in Statistics: Theory and Methods. 1997; 26:839–851.