CIENCIA 9(4), 448-451, 2001
Maracaibo, Venezuela

Three snapshots on bivariate random variables

José Luis Palacios

Centrode Estadísticas y Software Matemático. Universidad Simón Bolívar, Apartado 890000. Caracas, Venezuela.

Recibido: 01-03-01 Aceptado: 03-11-01

Abstract

We take a look at three instances where the subtleties of bivariate random variables arise: a "different" proof of the fact that $EXY = EXEY$ when X and Y are independent random variables; an example where $E\left(\frac{X}{Y}\right) = \frac{EX}{EY}$ but $E\left(\frac{Y}{X}\right) \neq \frac{EY}{EX}$ and a slick proof of the fact that $P(W \geq k) \leq P(Z \geq k)$ when W and Z are either both binomial with parameters, respectively, (n, x) and (n, y) with $x \leq y$, or both hypergeometric with parameters, respectively, (N,n,m_1) and (N,n,m_2) with $m_1 \leq m_2$.

Key words: Binominal; hypergeometric; independence.

Tres instantáneas sobre variables aleatorias bivariadas

Resumen

Damos una ojeada a tres situaciones que ilustran las sutilezas de las variables aleatorias bivariadas; una prueba "diferente" del resultado $EXY = EXEY$ cuando X y Y son independientes; un ejemplo donde $E\left(\frac{X}{Y}\right) = \frac{EX}{EY}$ pero $E\left(\frac{Y}{X}\right) \neq \frac{EY}{EX}$ y una prueba elegante de que $P(W \geq k) \leq P(Z \geq k)$ cuando W y Z son , o bien binomiales con parámetros, respectivamente, (n, x) y (n, y) con $x \leq y$, o ambas hipergeométricas con parámetros, respectivamente (N,n,m_1) y (N,n,m_2) con $m_1 \leq m_2$.

Palabras clave: Binomial; hipergeométrica; independencia.

1. Introduction

Anyone who has taught or has been taught an introductory course in Probability knows that there is a qualitative jump in complexity when going from studying a single random variable to studying a bivariate pair of random variables. There is much more to know about the joint pair than the individual behavior of the two marginals. The purpose of this note is to take a look at three instances where the subtleties of bivariate random variables are manifest: a "different" proof of the fact that $EXY = EXEY$ when X and Y are independent (EZ denotes throughout the expectation of the random variable Z); an example where $E\left(\frac{X}{Y}\right) = \frac{EX}{EY}$, but $E\left(\frac{Y}{X}\right) \neq \frac{EY}{EX}$ and an example where proving a result for univariate random variables (monotonicity of the tail probabilities of binominal or hypergeometric random varia-

*Autor para la correspondencia. Telf: 906-3233-4. Fax: (0212) 906 3232.

bles with respect to their parameters) is easier when one considers an appropriate bivariate pair.

2. Another proof of $EXY = EXEY$ when X and Y are independent

Introductory courses in Measure-theoretical Probability usually give a proof of the result

$$EXY = EXEY \qquad [1]$$

for X, Y independent random variables in either of these two ways [1]:

(i) prove the result for simple random variables, apply the monotone convergence theorem to extend the result to non-negative random variables and use linearity (plus an easy argumente to justify that X, Y independent implies X^+, Y^+, etc., are also independent) to extend the result to arbitrary random variables.

(ii) express the integral in the space Ω as an integral in $\mathbb{R}^2$ with respect to the measure $\mu^2(dx, dy)$ in the Borel sets of $\mathbb{R}^2$ generated by the pair (X, Y). Then independence of X, Y translates as $\mu^2(dx, dy)$ being equal to the product measure $\mu(dx) \times \mu(dy)$ and Fubini's theorem ends the job.

We present an alternative third proof that is based on a generalization of another well-known result that usually finds its way into most textbooks, at least as an exercise:

$$EX = \int_0^\infty P(X > t)dt \qquad [2]$$

for any non-negative random variable X. The proof of (2) is based on either "first-simples-then-monotone-convergence-then-linearity" (2) or Fubini's theorem, so in the end our proof can be seen as a reshuffling of the others, but we think it has some advantages, discussed at the end of the section. Here is the main "new" result:

Theorem 2.1

If X, Y are non-negative random variables then

$$EXY = \int_0^\infty \int_0^\infty P(X > x, Y > y)\, dx dy \qquad [3]$$

Proof. Write the right hand side as

$$\int_0^\infty \int_0^\infty \int_\Omega 1(x, \infty)x(y, \infty)(X(\omega)), Y(\omega))\, dP(\omega)\, dx dy =$$

by Fubini's theorem

$$\int_\Omega \int_0^\infty \int_0^\infty 1(x, \infty)x(y, \infty)(X(\omega), Y(\omega))\, dx dy\, dP(\omega).$$

The inner double integral is very easy to evaluate: it is just the area of a rectangle with sides of lengths $X(\omega)$ and $Y(\omega)$. and so we are left with

$$\int_\Omega X(\omega)Y(\omega)dP(\omega),$$

which is the left hand side of [3].

Theorem 2.2

If X and Y are independent random variables then $EXY = EXEY$.

Proof. Assume first that X and Y are non-negative. Then by [3] we can write

$$EXY = \int_0^\infty \int_0^\infty P(X > x, Y > y) dx dy.$$

Since X and Y are independent, the integrand can be written as $P(X > x)P(Y > y)$ and thus the double (Riemann) integral is just the product

$$\int_0^\infty P(X > x)\, dx \int_0^\infty P(Y > y)\, dy = EXEY\,.$$

Now we can finish the proof for arbitrary random variables just as in the version (i) of the proof discussed above.

The advantage of this proof is that the condition of independence that is used is the most elementary one:

$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$, where A and B are very simple Borel sets, namely intervals. Also, the Riemann integral in [3] looks less intimidating than the integral with respect to $\mu^2\,(dx, dy)$:

3. An example where $E\left(\frac{X}{Y}\right) = \frac{EX}{EY}$ but $E\left(\frac{Y}{X}\right) \neq \frac{EY}{EX}$

The moment generating fuction (m.g.f.) of a random variable X defined as $M(t) = Ee^{tX}$ is a basic tool in Probability theory. Most introductory books (2), discuss at varying lengths the salient feature of m.f.g.'s, namely, the fact that the moments of the variable X can be found by differentiation of its m.g.f.:

$$EX^k = M^{(k)}(0) \qquad [4]$$

It is not mentioned in the textbooks, however, that it is also possible to compute the inverse moments of a nonnegative random variable by integrating its m.g.f. The work of Cressie *et al.* (3) seems to be the first reference mentioning this fact. The next two theorems overlap with their results:

Theorem 3.1

If $M(t)$ is the m.g.f. of the nonnegative random variable X then

$$EX^{-k} = \int_{-\infty}^{0}\int_{-\infty}^{t_1}\ldots\int_{-\infty}^{t_{k-1}} M(t_k)dt_k \ldots dt_1, \qquad k \geq 1. \quad [5]$$

Proof. Only for $k = 1$. The generalizations are straightforward.

$$\int_{-\infty}^{0} M(t) = \int_{-\infty}^{o} Ee^{tX} = \int_{-\infty}^{0}\int_{\mathbb{R}^+} e^{tx} dF(x)dt,$$

where F is the distribution function of X. By the non-negativity of the integrand, we can apply Fubini's theorem and interchange the integrals to get:

$$\int_{\mathbb{R}^+}\left[\int_{-\infty}^{0} e^{tx}\,dt\right]dF(x) = \int_{\mathbb{R}^+}\frac{1}{x}dF(x) = E\left(\frac{1}{X}\right).$$

(Note that the non-negativity of the x variable is crucial in evaluating the inner integral).

If we combine differentiation and integration, we can prove results like the following formula for the expectation of the quotient of two random variables:

Theorem 3.2

Let $M(t_1, t_2) = Ee^{t_1X + T_2Y}$ be the joint m.g.f. of the pair (X, Y), and let X and Y be nonnegative with $EX < \infty$. Then

$$E\left(\frac{X}{Y}\right) = \int_{-\infty}^{0}\left[\frac{\partial}{\partial t_1} M(t_1, t_2)\Big|_{t_1=0}\right]dt_2. \qquad [6]$$

Proof. All we have to do is check that in the expression

$$\int_{-\infty}^{0}\left[\frac{\partial}{\partial t_1}\int_{\mathbb{R}^+ \times \mathbb{R}^+} e^{t_1x + t_2y}\; dF(x, y)\Big|_{t_1=0}\right]dt_2$$

we are allowed to (i) take the partial derivative inside the innermost integral and (ii) exchange the order of integration. That we may do (ii) follows as in the previous theorem; (i) is granted by the dominated convergence theorem and the fact that for $(t_1, t_2) \in \mathbb{R}^- \times \mathbb{R}^-$ and $(x, y) \in \mathbb{R}^+ \times \mathbb{R}^+$ we have

$$\left|\frac{\partial}{\partial t_1} e^{t_1x + t_2y}\right| \leq x,$$

and the bound is a dF-integrable function.

Example. Let us consider the pair (X, Y) with joint density function

$$f(x, y) = \frac{1}{\beta^4} x(y - x)e^{-\frac{y}{\beta}}, \quad 0 \leq x \leq y.$$

It is easy to see that X and Y are nonindependent variables with distributions Gamma $(2, \beta)$ and Gamma $(4, \beta)$ respectively. The joint m.g.f. is easily computed yielding

$$M(t_1, t_2) = [1 - \beta(t_1 + t_2)]^{-2}(1 - \beta t_2)^{-2}$$
$$\text{for } \frac{1}{\beta} \geq \max\{t_2, t_1 + t_2\}.$$

Now [6] yields

$$E\left(\frac{X}{Y}\right) = \int_{-\infty}^{0} \frac{\partial}{\partial t_1}[1 - \beta(t_1, t_2)]^{-2}(1 - \beta t_2)^{-2}\Big|_{t_1=0}\, dt_2$$
$$= 2\beta\int_{-\infty}^{0}(1 - \beta t_2)^{-5}\, dt_2 = \frac{1}{2},$$

and also

$$E\left(\frac{Y}{X}\right) = \int_{-\infty}^{0} \frac{\partial}{\partial t_2}[1 - \beta(t_1, t_2)]^{-2}(1 - \beta t_2)^{-2}\Big|_{t_2=0}\, dt_1$$
$$= 2\beta\int_{-\infty}^{0}\left[(1 - \beta t_1)^{-3} + (1 - \beta t_1)^{-2}\right]dt_1 = 3.$$

Notice that since $EX = 2\beta$ and $EY = 4\beta$, we have provided an example where $E\left(\frac{X}{Y}\right) = \frac{EX}{EY}$, but $E\left(\frac{Y}{X}\right) > \frac{EY}{EX}$.

When X and Y are independent, of course $E\left(\frac{X}{Y}\right) = EXE\left(\frac{1}{Y}\right)$, but even in this case there is no guarantee that the expectation of the quotient is the quotient of the expectations, as is remarked in example 4.13 of reference (4).

4. Univariate results with bivariate tricks

Perhaps the best known result of univariate results with bivariate tricks is the proof via couplings of the convergence of an ergodic Markov chain to its stationary distribution. We refer the reader to the monograph (2), theorem 8.6, for details. Here we illustrate the principle with a much more modest goal: consider W and Z to be binominal random variables with parameters, respectively (n, x) and (n, y), and $x \leq y$. We want to prove that $P(W \geq k) \leq P(Z \geq k)$ for any $k \geq 0$. One way to proceed is to compute directly both probabilities involved. Another more elegant way is to consider n independent copies of the bivariate 0-1-valued variables $(X_i, Y_i), 1 \leq i \leq n$, with joint distributions dictated by the probabilities $P(X_i = Y_i = 1) = x, P(X_i = 1, Y_i = 0) = 0$ and $P(X_i = Y_i = 0) = 1 - y$. If we define $W = \sum_{i=1}^{n} X_i$ and $Z = \sum_{i=1}^{n} Y_i$, then the distributions of W and Z are the binomials under consideration, and by construction

$$\{W \geq k\} \subset \{Z \geq k\}, \qquad [7]$$

so that

$$P(W \geq k) \leq P(Z \geq k). \qquad [8]$$

A similar idea can be applied in the case of the hypergeometric distribution. Let N, m_1, m_2 be integers with $N \geq m_2 \geq m_1$; we are going to define the joint distribution of a pair (W, Z) of variables in a descriptive way: assume that a box contains N tickets with two slots numbered 1 and 2, such that there are m_1 tickets with a white dot on both slots 1 and 2, $m_2 - m_1$ tickets with a black dot on slot 1 and a white dot on slot 2, and $N - m_2$ tickets with a black dot on both slots. Now take a sample of size n without replacement from the box and let W and Z be the number of white dots of the sampled tickets in, respectively, slots 1 and 2. Clearly W and Z are hypergeometric random variables with parameters, respectively (N, n, m_1) and (N, n, m_2). Again, by constructions, [7] holds and therefore the conclusion [8] also holds.

References

1. CHUNG K.L. A course in Probability theory, second edition, Academic Press, New York (USA), 1974.
2. BILLINGSLEY P. Probability and Measure, John Wiley and Sons, New York (USA), 1979.
3. CRESSIE N., DAVIS A.S., FOLKS J.L., POLICELLO G.E. The American Statistician 35: 148-150, 1981.
4. ROMANO J.P., SIEGEL A.F. Counterexamples in Probability and Statistics, Chapman & Hall, New York (USA), 1986.