Bivariate Rayleigh Distribution and its Properties

Ahmad Saeed Akhter
Institute of Statistics
University of the Punjab
Q.A. Campus, Lahore

Abdul Samad Hirai
Ex-Director
Institute of Statistics
University of the Punjab
Q.A. Campus, Lahore

1. Introduction

Rayleigh (1880) observed that the sea waves follow no law because of the complexities of the sea, but it has been seen that the probability distributions of wave heights, wave length, wave induce pitch, wave and heave motions of the ships follow the Rayleigh distribution. At present, several different quantities are in use for describing the state of the sea; for example, the mean height of the waves, the root mean square height, the height of the “significant waves” (the mean height of the highest one-third of all the waves) the maximum height over a given interval of the time, and so on. At present, the ship building industry knows less than any other construction industry about the service conditions under which it must operate. Only small efforts have been made to establish the stresses and motions and to incorporate the result of such studies in to design. This is due to the complexity of the problem caused by the extensive variability of the sea and the corresponding response of the ships. Although the problem appears feasible, yet it is possible to predict service conditions for ships in an orderly and relatively simple manner Rayleigh (1980) derived it from the amplitude of sound resulting from many independent sources. This distribution is also connected with one or two dimensions and is sometimes referred to as “random walk” frequency distribution. The Rayleigh distribution can be derived from the bivariate normal distribution when the variate are independent and random with equal variances. We try to construct bivariate Rayleigh distribution with marginal Rayleigh distribution function and discuss its fundamental properties.

2. Construction of Bivariate Rayleigh Distribution

a) $P(X \le x, Y \le y) = F(x, y)$ is defined as the joint distribution function of continuous random variables X and Y.
The distribution of X is

$$P(X \le x) = F_1(x) = F(x, \infty)$$

The distribution of Y is

$$P(Y \le y) = F_2(y) = F(\infty, y)$$

If $F(x, y)$ is continuous every where then we have

$$\frac{\partial^2 F}{\partial x \partial y} = f(x,y) \geq 0 \quad \text{and} \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dxdy = 1 \tag{1}$$

We denote $f_1(x)$ and $f_2(y)$ as the marginal density function of X and Y and $f(x/y)$ and $f(y/x)$ as the conditional density functions of X given Y and Y given X respectively

The conditional expectation of X given Y is

$E(x/y)$ and is called the regression function of x on y and similarly $E(y/x)$ is called the regression function of y on x.

$$E(x,y) = \int_{-\infty}^{\infty} yE(x/y) f_2(y)dy \tag{2}$$

The correlation ratio (squared) of x on y is defined as

$$\mu^2(x/y) = \frac{1}{\sigma_y^2}\int_{-\infty}^{\infty}\left[E(x) - E(x/y)\right]^2 f_2(y)\,dy \tag{3}$$

and similarly

$$\mu^2(y/x) = \frac{1}{\sigma_x^2}\int_{-\infty}^{\infty}\left[E(y) - E(y/x)\right]^2 f_1(x)\,dx \tag{4}$$

The correlation ratios are used as the index from the linearity of regression.

b) Of all probability distributions in two or more than the two variates the normal distribution is most important. The normal bivariate distribution have been studied from the times of Bravais and Karl Pearson because of its fundamental properties

i) The curves of equal probability density are concentric ellipses.

ii) The regression curves are straight lines which intersect at the expectations.

iii) With increasing values of one variable the conditional expectation of the other increases or decreases without limit.

iv) If $\rho = 0$ (the correlation coefficient) then the variates are independent in the probability sense.

v) If $\rho = 1$, an exact linear functional relationship exist between x and y.

3. Construction of Bivariate Rayleigh Distribution

As the Rayleigh distribution serves as an important model in noise theory, height of the sea waves, weapon testing and flight testing, it is reasonable to construct Bivariate Rayleigh distribution when the marginal distributions are Rayleigh. The bivariate distribution is not constructed by the marginal distribution. If the variates are independent then the bivariate distribution is determined by the marginal distributions. It has been observed that infinitely many bivariate distributions exist for given marginal distributions.

c) It has been observed by Gumbel (1960) that given two probability functions $F_1(x)$ and $F_2(y)$ a bivariate function $F(x, y)$ can be constructed by means of the equation

$$F(x, y) = F_1(x)F_2(y)[1+\alpha\{(1-F_1(x))(1-F_2(y)\}] -1 \le \alpha \le 1 \qquad (5)$$

Therefore, we have

$$\frac{\partial^2 F(x, y)}{\partial x \partial y} = f(x, y)$$

$$= f_1(x) f_2(y)\left[1+\alpha\{(2F_1(x)-1\}2F_2(y)-1\right] \qquad (6)$$

Of course when $\alpha = 0$ the random variables are independent

Let $F_1(x)$ and $F_2(y)$ be Rayleigh distribution function such that

$$F_1(x) = 1 - e^{\frac{-x^2}{-1^2}}$$

and $$F_2(y) = 1 - e^{\frac{-y^2}{-\lambda^2}} \qquad (7)$$

Using (2) we construct bivariate Rayleigh distribution as

$$f(x, y) = \frac{4xy}{\lambda^2}\left[e^{\frac{-(x^2+y^2)}{\lambda^2}} + \alpha\, e^{-\left(\frac{x^2+y^2}{\lambda^2}\right)} - 2\alpha\, e^{-\left(\frac{2x^2+y^2}{\lambda^2}\right)} - 2\alpha\, e^{-\left(\frac{x^2+2y^2}{\lambda^2}\right)} + 4\alpha\, e^{-2\left(\frac{x^2+y^2}{\lambda^2}\right)}\right. -1 \le \alpha \le 1$$

We note that

$$f(x, y) \ge 0 \text{ for all x and y} \qquad (8)$$

$$f(x, \infty) = f(\infty, y) = 0$$

$$f(u, o) = 0$$

Thus (3) is a proper bivariate density function with marginal distributions as Rayleigh.

4. Properties of the Bivariate Rayleigh Distribution

The marginal distribution of x given y is

$$f(x/y) = \frac{2x}{\lambda^2} e^{\frac{-x^2}{\lambda^2}}\left[1+\alpha - 2\alpha\, e^{\frac{-x^2}{\lambda^2}} - 2\alpha\, e^{\frac{-y^2}{\lambda^2}} + 4\alpha\, e^{\frac{-(x^2+y^2)}{\lambda^2}}\right] \qquad (9)$$

By symmetry we can find $f(y/n)$ by replacing y with x in (4).

The conditional expectation of X on Y is given by

$$g(x/y) = \frac{2}{\lambda^2}\int_0^\infty x^2\, e^{\frac{-x^2}{y^2}}\left[1+\alpha - 2\alpha\, e^{\frac{-x^2}{\lambda^2}} - 2\alpha\, e^{\frac{-y^2}{\lambda^2}} + 4\alpha\, e^{\frac{-(x^2+y^2)}{\lambda^2}}\right] xy \qquad (10)$$

On evaluating (5) we get

$$E(x/y) = \alpha_1 + \beta_1\, e^{\frac{-y^2}{\lambda^2}} \qquad (11)$$

where $\alpha_1 = \frac{\sqrt{\Pi}\alpha}{2}\left(1+\alpha e^{\frac{-\alpha}{\sqrt{2}}}\right)$:

$$\beta_1 = \left(\frac{1-\sqrt{2}}{\sqrt{2}}\right)\sqrt{\Pi}\alpha\lambda$$

Similarly by symmetry we get the conditional expectation of Y on X is

$$E(Y/X) = \alpha_1\beta_1 e^{\frac{-x^2}{\lambda^2}} \tag{12}$$

The regression curves are not linear but are exponential functions. The conditional expectation $E(X/Y)$ increases when y increases for $\alpha > 0$ and decreases for $\alpha < 0$.

Also we have

$$E(XY) = \int_0^\infty y\left(\alpha_1 + \beta_1 e^{\frac{-y^2}{\lambda^2}}\right)\frac{2y}{\lambda^2}dy$$

$$= \frac{\alpha\sqrt{\Pi}\alpha}{2} + \frac{\beta\sqrt{\Pi}\alpha}{4\sqrt{2}} \tag{13}$$

Thus correlation coefficient P is given in terms of the parameter α as

$$P = \frac{E(XY) - E(X)E(Y)}{\sqrt{Var(X)\,Var(Y)}} = 0.314\alpha \tag{14}$$

Where $E(X) = E(Y) = 0.8862\lambda$

$$Var(X) = Var(Y) = 0.2146\lambda^2 \tag{15}$$

Thus correlation varies in the narrow range $-0.314 \le P \le 0.314$

The coefficient of correlation is +ve or –ve as α is +ve or –ve respectively when $\alpha = 0$ we have $\beta = 0$ and the random variables are uncorrelated in the statistical sense.

The correlation ratio (squared) of x on y is from (3)

i.e. $\mu^2(x/y) = 03235\lambda$ (16)

and this measures the departure of linearity from the regression lines.

References

1. Gumbel, N.J. (1960). “Bivariate exponential distributions”, J. Amer Stat Association, 55, pp 698–707.
2. Rayleigh, L. (1880). “On the resultant of a large number vibrations of the same pitch and of arbitrary phase”, Phil, Mag 10, pp 73–80.