Clarkson University

Conditional Distributions and Densities

Definition: Conditional distribution of random variable Y given event m is defined as

$$F_Y(y \mid m) = P\{Y \le y \mid m\} = \frac{P\{Y \le y \cap m\}}{P(m)}.$$

Suppose $m = \{X \le x\}$, then

$$F_Y(y \mid X \le x) = \frac{P\{X \le x \cap Y \le y\}}{P\{X \le x\}} = \frac{F_{XY}(x, y)}{F_X(x)},$$

and

$$f_Y(y \mid X \le x) = \frac{\dfrac{\partial F_{XY}(x, y)}{\partial y}}{F_X(x)} = \frac{1}{F_X(x)} \int_{-\infty}^{x} f_{XY}(x_1, y) dx_1 .$$

Similarly,

$$F_Y(y \mid x_1 < X \le x_2) = \frac{\int_{x_1}^{x_2} f_{XY}(x, y) dx}{F_X(x_2) - F_X(x_1)}.$$

Conditional Distribution and Density of Y Given that X = x

Noting that

$$F_Y(y \mid X = x) = \lim_{\Delta x \to 0} F_Y(y \mid x < X \le x + \Delta x),$$

it follows that

$$F_Y(y \mid x = x) = \lim_{\Delta x \to 0} \frac{F_{XY}(x + \Delta x, y) - F_{XY}(x, y)}{F_X(x + \Delta x) - F_X(x)} = \frac{\dfrac{\partial F_{XY}(x, y)}{\partial x}}{\dfrac{dF_X(x)}{dx}} .$$

That is

$$F_Y(y \mid x = x) = \frac{\int_{-\infty}^{y} f_{XY}(x, y_1) dy_1}{f_X(x)},$$

and

$$f_Y(y \mid x) = f_Y(y \mid x = x) = \frac{f_{XY}(x, y)}{f_X(x)}.$$

Similarly one finds

$$F_X(x \mid y) = \frac{\int_{-\infty}^{x} f_{XY}(x_1, y) dx_1}{f_Y(y)},$$

and

$$f_X(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)}.$$

Conditional Expected Value

Definition: Conditional expected value of a function of a random variable is defined as

$$E\{g(Y) \mid m\} = \int_{-\infty}^{+\infty} g(y) f_Y(y \mid m) dy .$$

Conditional expected value of a function of a random variable given $X = x$ is defined as

$$E\{g(Y) \mid X = x\} = \int_{-\infty}^{+\infty} g(y) f_Y(y \mid x = x) dy .$$

That is,

$$E\{g(Y) \mid X = x\} = \frac{1}{f_X(x)} \int_{-\infty}^{+\infty} g(y) f_{XY}(x, y) dy$$

Chapman-Kolmogorov Equation

Noting that

$$f_X(x \mid y) = \frac{f_{XY}(x, y)}{f_Y(y)},$$

it follows that

$$f_X(x \mid y, z) = \frac{f_{XY}(x, y \mid z)}{f_Y(y \mid z)},$$

or

$$f_{XY}(x, y \mid z) = f_X(x \mid y, z) f_Y(y \mid z).$$

Integrating over y, we find

$$f_X(x \mid z) = \int_{-\infty}^{+\infty} f_X(x \mid y, z) f_Y(y \mid z) dy .$$

For Markov processes $f_X(x \mid y, z) = f_X(x \mid y)$. Hence,

$$f_X(x \mid z) = \int_{-\infty}^{+\infty} f_X(x \mid y) f_Y(y \mid z) dy .$$

This integral equation is the Chapman-Kolmogorov equation for a Markov process. It is a nonlinear equation for the (transition) conditional density function.

Clarkson University

Sample Mean and Sample Variance

Definition: Sample Mean

$$\overline{X} = \frac{X_1 + X_2 + \dots + X_n}{n}.$$

Definition: Sample Variance

$$\overline{V} = \frac{\left(X_1 - \overline{X}\right)^2 + \dots + \left(X_n - \overline{X}\right)^2}{n}.$$

Clearly $\overline{X}$ and $\overline{V}$ are random variables.

Consider the case that X_i have the same mean and variance and they form a sequence of uncorrelated random variables. It may be shown that

$$E\{\overline{X}\} = \eta, \ \sigma_{\overline{X}}^2 = \frac{\sigma^2}{n}, \ E\{\overline{V}\} = \frac{n-1}{n}\sigma^2.$$

When X_i are jointly normal with

$$f(x_1, \dots x_n) = \frac{1}{(2\pi)^{\frac{n}{2}}\sigma^n}\exp\left\{-\frac{x_1^2 + x_2^2 + \dots + x_n^2}{2\sigma^2}\right\},$$

the density functions of sample mean $\overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$ and sample variance $\overline{V} = \frac{1}{n}\sum_j \left(X_j - \overline{X}\right)^2$ are given as

$$f_{\overline{X}}(x) = \frac{1}{\sqrt{\frac{2\pi\sigma^2}{n}}} exp\left\{-\frac{nx^2}{2\sigma^2}\right\},$$

and

$$f_{\overline{V}}(v) = \frac{1}{2^{\frac{(n-1)}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)^{n-1}\Gamma\left(\frac{n-1}{2}\right)} v^{\frac{(n-3)}{2}} e^{-\frac{nv}{2\sigma^2}} U(v).$$

In statistics, chi-statistics, and chi-square statistics are used frequently. These are defined as

$$\chi = \sqrt{\sum_{j=1}^{n} X_j^2}\ ,\ \chi^2 = Y = \sum_{j=1}^{n} X_j^2\ .$$

The density functions of χ and $\chi^2 = Y$ are given as

$$f_X(\chi) = \frac{2}{2^{\frac{n}{2}}\sigma^n \Gamma\left(\frac{n}{2}\right)} \chi^{n-1} e^{-\frac{\chi^2}{2\sigma^2}} U(\chi),$$

and

$$f_{X^2}(y) = \frac{1}{2^{\frac{n}{2}}\sigma^n \Gamma\left(\frac{n}{2}\right)} y^{\frac{(n-2)}{2}} e^{-\frac{y}{2\sigma^2}} U(y).$$

Estimating Mean and Variance for Random Data

The mean and variance of a set of random data are given as

Mean: $\overline{X} = \frac{\sum_i X_i}{n}$,

Variance: $S^2 = \frac{\sum_i \left(X_i - \overline{X}\right)^2}{n-1}$.

If the random variables X_i have η and σ^2 as mean and variance then

$$E\{\overline{X}\} = \eta \quad \text{and} \quad \sigma_{\overline{X}}^2 = \frac{\sigma^2}{n}.$$

Theorem: If $\overline{X}$ is the mean of a random sample of size n taken from a population having the mean η and the variance σ^2, then

$$Z = \frac{\overline{X} - \eta}{\frac{\sigma}{\sqrt{n}}}$$

is a random variable whose distribution approaches that of the standard normal distribution as $n \to \infty$. i.e.,

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\}, \; P\{|Z| \le z^2\} = 2erf(z).$$

Note that

$$(P\{|Z| \le 1\} \approx 0.68, \; P\{|Z| \le 2\} = 0.85, \; P\{|Z| \le 3\} = 0.997).$$

Size of a Sample for a Required Accuracy

Let

error $= E = \left|\overline{X} - \eta\right|$,

and the set

$$|Z| \le z, \qquad \text{lead to} \qquad E \le z\frac{\sigma}{\sqrt{n}}.$$

The size of the sample needed is given by

$$n = \frac{z^2\sigma^2}{E^2}.$$

That is if the sample size is given by $n = \frac{z^2\sigma^2}{E^2}$, then with probability of $2erf(z)$ the error will not be more than E.

Example: Let $z = 3$, $\sigma = 2$, and $E = 0.01$. Then, $n = \frac{(9)(4)}{10^{-4}} = 36 \times 10^4$ data points are needed to estimate the mean with a probability of 0.997 and error less than 0.01.

For $E = 0.1$ under the same condition $n = 3600$.

Alternative Definition for Probability Density Function

The probability density function of a random variable $X(\xi)$ may be defined as

$$f_X(x) = E\{\delta(X-x)\}. \text{ (Stratonovich)}$$

This definition is equivalent to the common definition of the density function and the expected value. i.e.,

$$E\{\delta(X-x)\} = \int_{-\infty}^{+\infty} \delta(x_1 - x) f_X(x_1) dx_1 = f_X(x).$$

$$\begin{aligned} E\{g(x)\} &= E\left\{\int_{-\infty}^{+\infty} g(x)\delta(x-X) f_X(x) dx\right\} \\ &= \int_{-\infty}^{+\infty} g(x) E\{\delta(x-X)\} dx \\ &= \int_{-\infty}^{+\infty} g(x) f_X(x) dx \end{aligned} .$$

The entire theory of probability may be developed based on the alternative (Stratonovich) definition of the probability density function (pdf). For example, if $Y = g(X)$, then

$$f_Y(y) = \int_{-\infty}^{+\infty} \delta[g(x) - y] f_X(x) dx.$$

Using the property that $\delta[g(x)-y] = \sum_j \frac{\delta(x - x_j)}{|g'(x_j)|}$, where x_j is the solution to $g(x) = y$, it follows that

$$f_Y(y) = \int_{-\infty}^{+\infty} \sum_j \delta(x - x_j) f_X(x) dx = \sum_j \frac{f_X(x_j)}{|g'(x_j)|}.$$