Probability 2 - Notes 7

Continuous Random Variables

Definition. A random variable X is said to be a continuous random variable if there is a function $f_X(x)$ (the probability density function or p.d.f.) mapping the real line $\Re$ into $[0,\infty)$ such that for any open interval (a,b), $P(X \in (a,b)) = P(a < X < b) = \int_a^b f_X(x)dx$.

From the axioms of probability this gives:

(i) $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

(ii) The cumulative distribution function $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(u)du$. $F_X(x)$ is a monotone increasing function of x with $F_X(-\infty) = 0$ and $F_X(\infty) = 1$.

(iii) $P(X = x) = 0$ for all real x.

From calculus, $f_X(x) = \frac{dF_X(x)}{dx}$ for all points for which the p.d.f. is continuous and hence the c.d.f. is differentiable.

Expectations, Moments and the Moment Generating Functions

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx$$

The raw moments are moments about the origin. The r^{th} raw moment is $\mu_r = E[X^r]$. Note that μ_1 is just the mean μ.

The moment generating function (m.g.f.) $M_X(t) = E[e^{tX}]$.

For a discrete random variable $M_X(t) = G_X(e^t)$.

For a continuous random variable $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx$.

Properties of the M.G.F.

(i) If you expand $M_X(t)$ in a power series in t you obtain $M_X(t) = \sum_{r=0}^{\infty} \frac{\mu_r t^r}{r!}$. So the m.g.f. generates the raw moments.

(ii) $\mu_r = E[X^r] = M_X^{(r)}(0)$, where $M_X^{(r)}(t)$ denotes the r^{th} derivative of $M_X(t)$ with respect to t.

(iii) The m.g.f. determines the distribution.

Other properties (similar to those for the p.g.f.) will be considered later once we have looked at joint distributions.

Standard Continuous Distributions

Uniform Distribution. All intervals (within the support of the p.d.f.) of equal length have equal probability of occurrence. Arises in simulation. Simulated values $\{u_j\}$ from a uniform distribution on $(0,1)$ can be transformed to give simulated values $\{x_j\}$ of a continuous r.v. X with c.d.f. F by taking $x_j = F^{-1}(u_j)$.

$X \sim U(a,b)$ if

$$f_X(x) = \begin{cases} \frac{1}{(b-a)} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

$E[X] = \frac{a+b}{2}$ and $Var(X) = \frac{(b-a)^2}{12}$.

$M_X(t) = \frac{e^{bt} - e^{at}}{t(b-a)}$. This exists for all real t.

Exponential Distribution. Used for the time till the first event if events occur randomly and independently in time at constant rate. Used as a survival distribution for an item which remains as 'good as new' during its lifetime.

$X \sim Exp(\theta)$ if

$$f_X(x) = \begin{cases} \theta e^{-\theta x} & \text{if } 0 < x < \infty \\ 0 & \text{otherwise} \end{cases}$$

$E[X] = \frac{1}{\theta}$ and $Var(X) = \frac{1}{\theta^2}$.

$M_X(t) = \left(1 - \frac{t}{\theta}\right)^{-1}$. This exists for $t < \theta$.

Gamma Distribution. Exponential is special case. Used as a survival distribution. When $\alpha = n$, gives the time until the n^{th} event when events occur randomly and independently in time.

$X \sim Gamma(\theta, \alpha)$ if

$$f_X(x) = \begin{cases} \frac{\theta^\alpha x^{\alpha-1} e^{-\theta x}}{\Gamma(\alpha)} & \text{if } 0 < x < \infty \\ 0 & \text{otherwise} \end{cases}$$

The Gamma function is defined for $\alpha > 0$ by $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$. It is then simple to show that the p.d.f. integrates to one by making a simple change of variable ($y = \theta x$) in the integral.

It is easily shown using integration by parts that $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$. Therefore when n is a positive integer $\Gamma(n) = (n-1)!$.

$E[X] = \frac{\alpha}{\theta}$ and $Var(X) = \frac{\alpha}{\theta^2}$.

$M_X(t) = \left(1 - \frac{t}{\theta}\right)^{-\alpha}$. This exists for $t < \theta$.

Note: The Chi-squared distribution ($X \sim \chi^2_n$) is just the gamma distribution with $\theta = 1/2$ and $\alpha = n/2$. This is an important distribution in normal sampling theory.

Normal Distribution. Important in statistical modelling where normal error models are commonly used. It also serves as a large sample approximation to the distribution of efficient estimators in statistics.

$X \sim N(\mu, \sigma^2)$ if

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

To show that the p.d.f. integrates to 1 by a simple change of variable in the integral to $z = (x-\mu)/\sigma$ we just need to show that $\int_{-\infty}^{\infty} e^{-z^2/2} = \sqrt{2\pi}$. We show this at the end of Notes 5.

$E[X] = \mu$ and $Var(X) = \sigma^2$.

$M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$. This exists for all t.

Example deriving the m.g.f. and finding moments

$X \sim N(\mu, \sigma^2)$.

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)} dx$$

In the integral make the change of variable to $y = (x-\mu)/\sigma$. Then

$$M_X(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2 + t(\mu+\sigma y)} dy = e^{\mu t + \sigma^2 t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(y-\sigma t)^2/2} dy = e^{\mu t + \sigma^2 t^2/2}$$

Finding $E[X]$ and $E[X^2]$. Differentiating gives $M_X'(t) = (\mu + \sigma^2 t) e^{\mu t + \sigma^2 t^2/2}$ and

$$M_X^{(2)}(t) = (\sigma^2) e^{\mu t + \sigma^2 t^2/2} + (\mu + \sigma^2 t)^2 e^{\mu t + \sigma^2 t^2/2}$$

Therefore $E[X] = M_X'(0) = \mu$ and $E[X^2] = M_X^{(2)}(t) = \sigma^2 + \mu^2$.

Transformations of random variables.

Theorem. *Let the interval A be the support of the p.d.f. $f_X(x)$. If g is a 1:1 continuous map from A to an interval B with differentiable inverse, then the r.v. $Y = g(X)$ has p.d.f.*

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|$$

Proof. This is easily shown using equivalent events. The function $g(x)$ will either be (a) strictly monotone increasing; or (b) strictly monotone decreasing. We consider each case separately.

Case (a)

$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \le g^{-1}(y)) = F_X(g^{-1}(y))$$

Differentiating and noting that $\frac{dg^{-1}(y)}{dy} > 0$ gives

$$f_Y(y) = f_X(g^{-1}(y)) \times \frac{dg^{-1}(y)}{dy} = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|$$

Case (b)

$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \ge g^{-1}(y)) = 1 - F_X(g^{-1}(y))$$

Differentiating and noting that $\frac{dg^{-1}(y)}{dy} < 0$ gives

$$f_Y(y) = -f_X(g^{-1}(y)) \times \frac{dg^{-1}(y)}{dy} = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|. \quad \square$$

Example $X \sim N(\mu, \sigma^2)$. Let $Y = \frac{X-\mu}{\sigma}$. Then $g^{-1}(y) = \mu + \sigma y$. Therefore $\frac{dg^{-1}(y)}{dy} = \sigma$. Hence

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right| = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-y^2/2} \times \sigma = \frac{1}{\sqrt{2\pi}}e^{-y^2/2}$$

The support for the p.d.f. X, $(-\infty, \infty)$, is mapped onto $(-\infty, \infty)$ so this is the support of the p.d.f. for Y. Therefore $Y \sim N(0, 1)$.

Transformations which are not 1:1. You can still find the c.d.f. for the transformed variable by writing $F_Y(y)$ as an equivalent event in terms of X.

Example. $X \sim N(0, 1)$ and $Y = X^2$. The support for the p.d.f. of Y is $[0, \infty)$. For $y > 0$,

$$F_Y(y) = P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

Differentiating with respect to y gives, for $y > 0$,

$$f_Y(y) = f_X(\sqrt{y})\frac{1}{2\sqrt{y}} - f_X(-\sqrt{y})\frac{-1}{2\sqrt{y}} = \frac{y^{-1/2}e^{-y/2}}{2^{1/2}\sqrt{\pi}}$$

This is just the p.d.f. for a χ_1^2. Note that this implies that $\Gamma(1/2) = \sqrt{\pi}$ because the constant in the p.d.f. is determined by the function of y and the range (suppport of the p.d.f.) since the p.d.f. integrates to one.

Note for the normal p.d.f.

Let $A = \int_{-\infty}^{\infty} e^{-z^2/2}dz$. Note that $A > 0$. Then

$$A^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}dxdy$$

Making the change to polar co-ordinates (Calculus 2) gives

$$A^2 = \int_0^{2\pi}\int_0^{\infty} e^{-r^2/2}rdrd\theta = 2\pi$$

Hence $A = \sqrt{2\pi}$.