

## Functions of Random Variables

Let's consider a continuous r.v.,  $Y$ , which is a differentiable, increasing function of a second continuous r.v.,  $X$ :

$$Y = g(X).$$

Because  $g$  is differentiable and increasing,  $g'$  and  $g^{-1}$  are guaranteed to exist. Because  $g$  maps all  $x \leq s \leq x + \Delta x$  to  $y \leq t \leq y + \Delta y$ :

$$\int_x^{x+\Delta x} f_X(s) ds = \int_y^{y+\Delta y} f_Y(t) dt.$$

It follows that for small  $\Delta x$ :

$$f_Y(y)\Delta y \approx f_X(x)\Delta x.$$

## Functions of Random Variables (contd.)

Dividing by  $\Delta y$ , we get an approximate expression for  $f_Y$  in terms of  $f_X$ :

$$f_Y(y) \approx f_X(x) \frac{\Delta x}{\Delta y}.$$

## Functions of Random Variables (contd.)

This expression is exact in the limit  $\Delta x \rightarrow 0$ :

$$\begin{aligned} f_Y(y) &= \lim_{\Delta x \rightarrow 0} f_X(x) \frac{\Delta x}{\Delta y} \\ &= \lim_{\Delta x \rightarrow 0} f_X(x) \frac{1}{\Delta y / \Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{f_X(x)}{\Delta y / \Delta x}. \end{aligned}$$

From calculus we know that:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{g(x + \Delta x) - g(x)}{\Delta x} = g'(x).$$

Consequently

$$f_Y(y) = \frac{f_X(x)}{g'(x)}.$$

## Functions of Random Variables (contd.)

Substituting  $g^{-1}(y)$  for  $x$ :

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}$$

yields an expression for  $f_Y$  in terms of  $g'$ ,  $g^{-1}$ , and  $f_X$ .

## Linear Example

Consider a continuous r.v.,  $Y$ , which is a linear function of a continuous r.v.,  $X$ . Specifically,  $Y = aX + b$ . It follows that

$$\begin{aligned}g(x) &= ax + b \\g'(x) &= a \\g^{-1}(y) &= (y - b)/a.\end{aligned}$$

Substituting the above into

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{g'(g^{-1}(y))}$$

yields

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y - b}{a}\right).$$

## Linear Example (contd.)

Let  $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-(x-\mu)^2/2}$ , then

$$f_Y(y) = \frac{1}{a\sqrt{2\pi}}e^{-\left(\frac{y-b}{a}-\mu\right)^2/2}.$$

## Linear Example (contd.)

Unfortunately, there is a problem. When  $a = -1$ , the function  $g$  is not increasing (it is decreasing). Consequently,

$$-f_Y(y)\Delta y \approx f_X(x)\Delta x.$$

It follows that for decreasing functions,

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{-g'(g^{-1}(y))}.$$

However, we can derive a function which is correct in both cases by replacing  $g'(\cdot)$  with  $|g'(\cdot)|$  in the expression relating  $f_Y(\cdot)$  and  $f_X(\cdot)$ :

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|}.$$

## Quadratic Example

Consider a continuous r.v.,  $Y$ , which is a quadratic function of a continuous r.v.,  $X$ . Specifically,  $Y = X^2$ . It follows that

$$\begin{aligned}g(x) &= x^2 \\g'(x) &= 2x \\g^{-1}(y) &= \sqrt{y}.\end{aligned}$$

Substituting the above into

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|}$$

yields

$$f_Y(y) = \frac{f_X(\sqrt{y})}{|2\sqrt{y}|}.$$

## Quadratic Example (contd.)

Let  $f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ , then

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

Unfortunately, there is a problem:

$$\int_0^{\infty} f_Y(y)dy = \int_0^{\infty} \frac{e^{-y/2}}{\sqrt{2\pi}|2\sqrt{y}|}dy = \frac{1}{2} \neq 1.$$

Can anyone see the mistake?

## Quadratic Example (contd.)

The mistake is that two different values of  $Y$  satisfy  $Y = X^2$ :

$$g^{-1}(y) = \pm\sqrt{y}.$$

We decided to use the positive square root arbitrarily and ignored the negative square root. Hence the factor of two error. In general, if a function does not have a unique inverse, we must sum over all possible inverse values:

$$f_Y(y) = \sum_{i=1}^n \frac{f_X(g_i^{-1}(y))}{|g'(g_i^{-1}(y))|}.$$

## Quadratic Example (contd.)

Let  $g_1^{-1}(y) = \sqrt{y}$  and  $g_2^{-1}(y) = -\sqrt{y}$ ,  
then

$$\begin{aligned} f_Y(y) &= \frac{f_X(g_1^{-1}(y))}{|g'(g_1^{-1}(y))|} + \frac{f_X(g_2^{-1}(y))}{|g'(g_2^{-1}(y))|} \\ &= \frac{e^{-y/2}}{\sqrt{2\pi y}}. \end{aligned}$$

The above p.d.f. defines a distribution called the *chi square* distribution.

## Kinetic Energy

Recall from elementary physics, that the *kinetic energy*,  $K$ , of a moving particle is given by

$$K = \frac{1}{2}mV^2$$

where  $m$  is mass and  $V$  is velocity. Let  $V$  be a normally distributed random variable with mean,  $\mu$ , and variance,  $\sigma^2$ :

$$f_V(v) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(v-\mu)^2/2\sigma^2}.$$

We would like to compute  $f_K(k)$ , the p.d.f for the continuous random variable,  $K$ .

## Kinetic Energy (contd.)

We start by computing,  $g'$ ,  $g_1^{-1}$ , and  $g_2^{-1}$ :

$$\begin{aligned}g(v) &= \frac{1}{2}mv^2 \\g'(v) &= mv \\g_1^{-1}(k) &= \sqrt{2k/m} \\g_2^{-1}(k) &= -\sqrt{2k/m}.\end{aligned}$$

Substituting  $f_V(v)$  and the above into

$$f_K(k) = \frac{f_V(g_1^{-1}(k))}{|g'(g_1^{-1}(k))|} + \frac{f_V(g_2^{-1}(k))}{|g'(g_2^{-1}(k))|}$$

yields

$$f_K(k) = \frac{e^{-\left(\sqrt{2k/m}-\mu\right)^2/2\sigma^2} + e^{-\left(-\sqrt{2k/m}-\mu\right)^2/2\sigma^2}}{\sigma\sqrt{\pi km}}.$$