1.5 Variable Transformations

Let the RV X have pdf $f(x)$ and let $h(x)$ denote some function of X. (e.g. if x is the colour of a star, then $h(x)$ could be the temperature). $H = h(x)$ is itself a RV with pdf $g(h)$, say. How are $f(x)$ and $g(h)$ related? We have that

$$f(x)\,dx = \text{prob}(x < X < x + dx)$$

and we require to find $g(h)$ such that

$$g(h)\,dh = \text{prob}(h < H < h + dh)$$

Suppose first that $h(x)$ is one-to-one, i.e. x maps to a unique h and vice versa. Hence the inverse function $x = x(h)$ exists, and we can write f as a function of h, i.e.

$$f(x) \equiv f(x(h))$$

This function is *not* $g(h)$, however, since we also have to transform the infinitesimal dx (just as with changing variables in integration). Thus

$$dx = \left|\frac{dx}{dh}\right| dh$$

where the modulus is required because probability is never negative. Combining the above expressions:-

$$f(x)\,dx = f(x(h)) \left|\frac{dx}{dh}\right| dh = g(h)\,dh$$

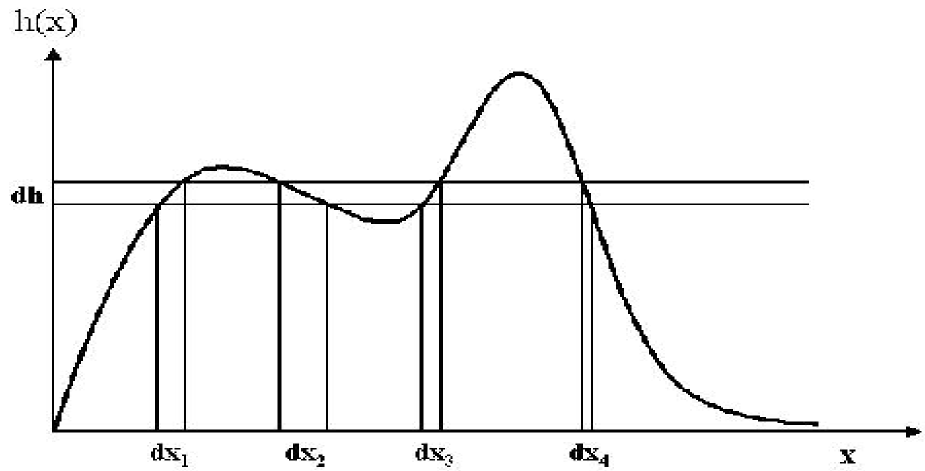
If $h(x)$ is *not* one-to-one then we must sum over all values of x for which $h(x) = h$, or more precisely over the small intervals, dx_i, corresponding to dh (see Figure 7). Thus

$$\text{prob}(h < H < h{+}dh) = \text{prob}(x_1 < x < x_1{+}dx_1){+}\text{prob}(x_2 < x < x_2{+}dx_2){+}\text{prob}(x_3 < x < x_3{+}dx_3){+}\ldots$$

It then follows that

$$g(h)\,dh = \sum_{h(x_i)=h} f(x_i(h)) \left|\frac{dx}{dh}\right|_{x_i(h)} dh$$

Figure 7: Variable transformation when h is not one-to-one



1.6 : Probability Integral Transform

One variable transformation merits special consideration. Suppose X has PDF $f(x)$ and CDF $F(x)$. Define $h(x) \equiv F(x)$, which is one-to-one. Then:-

$$\begin{aligned} g(h)\,dh &= f(x(h)) \left|\frac{dx}{dh}\right| dh \\ &= f(x(h)) \left|\frac{dh}{dx}\right|^{-1} dh \end{aligned} \tag{1}$$

Since $h(x) = F(x)$, $dh/dx = f(x)$. Thus

$$g(h)\,dh = \frac{f(x)}{f(x)}\,dh = 1.dh$$

i.e. the pdf of H is the uniform distribution, $U(0,1)$, (since $0 < F(x) < 1$).

This important result shows that we can always transform the PDF of any RV into the simple form of $U(0,1)$, **provided** we know the CDF of the original RV. This approach can be used in generating random numbers numerically (see e.g. Numerical Recipes, Chap 17.).

1.7 : Multivariate Distributions

Thus far we have considered only the properties of distributions of a single (univariate) RV. We now extend to the **multivariate** case of two or more RVs.

1.7.1 : Joint PDF

The **joint PDF** of two RVs, X_1 and X_2 is $p(x_1, x_2)$. Then,

$$\text{Prob}(a_1 < X_1 < b_1 \text{ and } a_2 < X_2 < b_2) \quad = \quad \int_{a_1}^{b_1} \int_{a_2}^{b_2} p(x_1, x_2)\, dx_1 dx_2$$

Extension to more than two RVs is carried out in the obvious way.

1.7.2 : Marginal Distributions

The **marginal PDF**, $p_1(x_1)$ of X_1 is defined by

$$p_1(x_1) \quad = \quad \int_{-\infty}^{\infty} p(x_1, x_2)\, dx_2$$

and is a PDF in the usual sense that

1. $p_1(x_1) \geq 0$, for all x_1
2. $\text{Prob}(a < X_1 < b) \quad = \quad \int_a^b p_1(x_1) dx_1$
3. $\int_{-\infty}^{\infty} p_1(x_1) dx_1 \quad = \quad 1$

Similarly, the marginal PDF of X_2 is

$$p_2(x_2) \quad = \quad \int_{-\infty}^{\infty} p(x_1, x_2)\, dx_1$$

In general, given any multivariate PDF, we may find the marginal PDF of any subset of the $X_1, ..., X_n$ by integrating over all other variables. e.g.

$$p_{13}(x_1, x_3) \quad = \quad \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1, ..., x_n)\, dx_2 dx_4 dx_5 ... dx_n$$

1.7.3 : Conditional Distributions

Consider the joint PDF, $p(x_1, x_2)$, of X_1 and X_2. Suppose we observe X_1 to have the value x_1, but do not observe X_2. We want a function that describes the PDF of X_2, given that $X_1 = x_1$ (usually simply stated as 'given x_1'). This function is known as the **conditional** PDF of X_2, written as $p(x_2|x_1)$, and defined by

$$p(x_2|x_1) = \frac{p(x_1, x_2)}{p_1(x_1)}$$

i.e. the conditional PDF is obtained by dividing the joint PDF of X_1 and X_2 by the marginal PDF of x_1 (provided $p_1(x_1) \neq 0$). Similarly

$$p(x_1|x_2) = \frac{p(x_1, x_2)}{p_2(x_2)}$$

Note that we can write

$$\begin{aligned} p(x_1, x_2) &= p(x_1|x_2)p_2(x_2) \\ &= p(x_2|x_1)p_1(x_1) \end{aligned}$$

This is known as **Bayes' formula**.

Extension to more than 2 RVs is again straightforward. For example,

$$p(x_1, x_3|x_2, x_4) = \frac{p(x_1, x_2, x_3, x_4)}{p_{24}(x_2, x_4)}$$

1.8 : Statistical Independence

If the conditional PDF of X_2 given x_1 does *not* depend on x_1, this means that X_1 and X_2 are statistically independent, since the observed value of X_2 is unaffected by the observed value of X_1.

Equivalently, X_1 and X_2 are independent if and only if the joint PDF of X_1 and X_2 can be written as the product of their marginal PDFs, i.e.

$$p(x_1, x_2) = p_1(x_1)\, p_2(x_2)$$

Again, we extend in the obvious way. The RVs $X_1, ..., X_n$ are **mutually independent** if and only if their joint PDF can be written as the product of their marginal pdfs. i.e.

$$p(x_1, x_2, ..., x_n) = p_1(x_1)p_2(x_2)...p_n(x_n)$$