PROBABILITY

QSNAKE EDITION

1. Joint Probability Distribution of Functions of Random Variables.

Recall. Change of Variables in a Double integral
Suppose that T is a C^1 transformation whose Jacobian is nonzero and that T maps a region S in the $uv-plane$ onto a region R in the xy-plane. Suppose that f is continuous on R and that R and S are type I or type II plane regions. Suppose also that T is one-to-one, except perhaps on the boundary of S. Then

$$\iint_R f(x,y)dA = \iint_S f(x(u,v),y(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right|$$

In Probability, we change the notation here:
X_1, X_2 be jointly continuous random variables with joint probability density function F_{X_1,X_2}. Suppose that $Y_1 = g_1(X_1,X_2)$ and $Y_2 = g_2(X_1,X_2)$ for some functions g_1 and g_2
we denote Jacobian as

$$J(x_1,x_2) = \begin{vmatrix} \dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2} \\ \dfrac{\partial g_2}{\partial x_1} & \dfrac{\partial g_2}{\partial x_2} \end{vmatrix} = \frac{\partial g_1}{\partial x_1}\frac{\partial g_2}{\partial x_2} - \frac{\partial g_1}{\partial x_2}\frac{\partial g_2}{\partial x_1} \neq 0$$

Example(p. 279)
Let X_1 and X_2 be jointly continuous random variables with probability density function f_{X_1,X_2}. Let $Y_1 = X_1 + X_2,\ Y_2 = X_1 - X_2$. Find the joint density function of Y_1 and Y_2 in terms of f_{X_1,X_2}

2. Little Note in Independent Random Variable.

Theorem. *Two random variables X_1 and X_2 with joint pdf $f(x_1, x_2)$ are independent if and only if :*

(1) the 'support set', $\{(x_1, x_2) \mid f(x_1, x_2) > 0\}$, is a Cartesian product, $A \times B$, and

(2) the joint pdf can be factored into the product of functions of x_1 and x_2, $f(x_1, x_2) = g(x_1)h(x_2)$

Example 1

The joint pdf of a pair X_1 and X_2 is

$$f(x_1, x_2) = 8x_1x_2 \ \ 0 < x_1 < x_2 < 1$$

and zero otherwise. This function can clearly be factored according to part (2) of the theorem, but the support set, $\{(x_1, x_2) \mid 0 < x_1 < x_2 < 1\}$, is a triangular region that cannot be represented as a Cartesian product. Thus, X_1 and X_2 are dependent.

Example 2

Consider now a pair X_1 and X_2 with joint pdf

$$f(x_1, x_2) = x_1 + x_2 \ \ 0 < x_1 < 1, \ 0 < x_2 < 1$$

and zero otherwise. In this case the support set is $\{(x_1, x_2) \mid 0 < x_1 < 1 \text{ and } 0 < x_1 < 1\}$, which can be represented as $A \times B$, where A and B are both the open interval $(0, 1)$. However, part (2) of the theorem is not satisfied because $x_1 + x_2$ cannot be factored as $g(x_1)h(x_2)$. Thus X_1 and X_2 are dependent.

3. Correlation.

Definition. *If X and Y are random variables with variances σ_X^2 and σ_Y^2 and covariance $\sigma_{XY} = Cov(X, Y)$, then the correlation coefficient of X and Y is*

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

The random variables X and Y are said to be uncorrelated if $\rho = 0$; otherwise they are said to be correlated.

Theorem. *If ρ is the correlation coefficient of X and Y, then*

$$-1 \leq \rho \leq 1$$

and $\rho = \pm 1$ if and only if $Y = aX + b$ with probability 1 for some$a \neq 0$ and b

Theorem. *If $E(Y \mid x)$ is a linear function of x, then*

$$E(Y \mid x) = \mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x - \mu_1)$$

and

$$E_X[Var(Y \mid X)] = \sigma_2^2(1 - \rho^2)$$

Proof. If $E(Y \mid x) = ax + b$, then

$$\mu_2 = E(Y) = E_X[E(Y \mid X)] = E_X(aX + b) = a\mu_1 + b$$

and

$$\begin{aligned}
\sigma_{XY} &= E[(X - \mu_1)(Y - \mu_2)] \\
&= E[(X - \mu_1)Y] \\
&= E_X\{E[(X - \mu_1)Y \mid X]\} \\
&= E_X[(X - \mu_1)E(Y \mid X)] \\
&= E_X[(X - \mu_1)(aX + b)] \\
&= a\sigma_1^2
\end{aligned}$$

In second line, we use the fact:$Cov(X,Y) = E[XY] - E[X]E[Y]$
in last 2 line, we use the fact:$E[g(X)Y \mid x] = g(x)E[Y \mid X]$
Thus,

$$a = \frac{\sigma_{XY}}{\sigma_1^2} = \rho\frac{\sigma_2}{\sigma_1} \text{ and } b = \mu_2 - \rho\frac{\sigma_2}{\sigma_1}\mu_1$$

■

Proof.

$$\begin{aligned} E_X[Var(Y \mid X)] &= Var(Y) - Var_X\left[\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X - \mu_1)\right] \\ &= Var(Y) - \rho^2\frac{\sigma_2^2\sigma_1^2}{\sigma_1^2} \\ &= \sigma_2^2(1 - \rho^2) \end{aligned}$$

■

Limit Theorem.

Proposition (Markov's inequality)**.** *If X is a random variable that takes only nonnegative values, then for any value $a > 0$,*

$$P\{X \geq a\} \leq \frac{E[X]}{a}$$

Proof. For $a > 0$, let

$$I = \begin{cases} 1 & \text{if } X \geq a \\ 0 & \text{otherwise} \end{cases}$$

and not that, since $X \geq 0$

$$I \leq \frac{X}{a}$$

Taking expectations of the preceding inequality yields

$$E[I] \leq \frac{E[X]}{a}$$

which, because $E[I] = P\{X \geq a\}$, proves the result ∎

Proposition (Chebyshev's inequality)**.** *If X is a random variable with finite mean μ and variance σ^2, then for any value $k > 0$*

$$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

Proof. Since $(X - \mu)^2$ is a nonnegative random variable, we can apply Markov's inequality (with $a = k^2$) to obtain

$$P\{(X - \mu)^2 \geq k^2\} \leq \frac{E[(X - \mu)^2]}{k^2}$$

But since $(X - \mu)^2 \geq k$ if and only if $|X - \mu| \geq k$

$$P\{X - \mu \geq k\} \leq \frac{E[(X - \mu)^2]}{k^2} = \frac{\sigma^2}{k^2}$$

∎

Proposition. *If* $Var(X) = 0$, *then*

$$P\{X = E[X]\} = 1$$

In other words, the only random variables having variances equal to 0 *are those that are constant with probability* 1.

Proof. By Chebyshev's inequality, we have, for any $n \geq 1$,

$$P\left\{|X - \mu| > \frac{1}{n}\right\} = 0$$

Letting $n \to \infty$ and using the continuity property of probability yields

$$0 = \lim_{n\to\infty} P\left\{|X - \mu| > \frac{1}{n}\right\} = P\left\{. \lim_{n\to\infty} \left\{|X - \mu| > \frac{1}{n}\right\}\right\} = P\{X \neq \mu\}$$

and the result is established ■