Section 6.6: Bivariate Transformations using Jacobians

Often, we are interested in find the probability distribution of a function of two or more random variables and/or the joint distribution of functions of multivariate random variables.

For example, suppose two random variables Y_1 and Y_2 have joint pdf/pmf $f_{Y_1,Y_2}(y_1, y_2)$.

We are sometimes interested in a new bivariate random vector (U, V) defined by $U = g_1(Y_1, Y_2)$ and $V = g_2(Y_1, Y_2)$, and we want to find either the CDF/pdf/pmf of U or V, or the joint pdf/pmf of (U, V).

In such situations, we can extend the methods used in the univariate case to the bivariate (or multivariate) case.

Example:

A shipping company handles containers in three different sizes:
(1) 27 ft^3, (2) 125 ft^3, and (3) 512 ft^3.
Let Y_i $(i = 1, 2, 3)$ denote the number of type i containers shipped during a given week.

Total volume shipped =

Example:

A gas station sells three grades of gasoline: regular, extra, and super. These are priced at \$3.00, \$3.20, and \$3.40 per gallon, respectively. Let Y_1, Y_2, and Y_3 denote the amounts of these grades purchased (gallons) on a particular day.
Suppose the Y_i are independent and each Y_i has pdf of $f(y_i)$.

The revenue from sales is

Find the probability that revenue exceeds 4500. That is, find

Example:

Five automobiles of the same type are to be driven on a 300-mile trip. The first two will use an economy brand of gasoline, and the other three will use a name brand.
Let Y_1, Y_2, Y_3, Y_4, and Y_5 be the observed fuel efficiencies (mpg) for the five cars.

Suppose these variables are independent and normally distributed with
$\mu_1 = \mu_2 = 20$, $\mu_3 = \mu_4 = \mu_5 = 21$, $\sigma_1 = \sigma_2 = 2$ and $\sigma_3 = \sigma_4 = \sigma_5 = 1.1$.
U is a measure of the difference in efficiency between economy gas and name-brand gas and defined as

$$U = \frac{Y_1 + Y_2}{2} - \frac{Y_3 + Y_4 + Y_5}{3}$$

Compute $P(U \leq 0)$ and $P(-1 \leq U \leq 1)$.

Example:

Suppose your waiting time for a bus in the morning is denoted by Y_1 and uniformly distributed on [0, 8] minutes, whereas waiting time in the evening is denoted by Y_2 and uniformly distributed on [0, 10] minutes independent of morning waiting time.

a. If you take the bus each morning and evening for a week, what is your total expected waiting time?

b. What is the variance of your total waiting time per week?

c. What is the probability that the difference between morning and evening waiting times on a given day is less than 5 minutes?

Now, suppose Y_1 and Y_2 are (absolutely) **continuous** random variables. If there is a one-to-one transformation from the support of $(Y_1,\ Y_2)$ to the support of (U_1, U_2), the **Jacobian** method discussed for the univariate case can be extended to find the joint distribution of (U_1, U_2) in the bivariate case.

Bivariate Transformations for <u>Continuous</u> Random Variables using Jacobian:

Suppose Y_1 and Y_2 are (absolutely) continuous random variables with joint pdf $f_{Y_1,Y_2}\ (y_1, y_2)$.

Let $U_1\ = h_1(Y_1, Y_2)$

Let $U_2\ =\ h_2(Y_1, Y_2)$

Suppose the transformation pair $u_1\ = h_1(y, y_2)$ and $u_2\ =\ h_2(y_1, y_2)$ is **one-to-one.**

Then, for each $(u_1,\ u_2)$ in the support of $(U_1,\ U_2)$,

$y_1 = h_1^{-1}(u_1, u_2)$

and $y_2 = h_2^{-1}(u_1, u_2)$

If y_1 and y_2 have continuous partial derivatives with respect to u_1 and u_2, and **absolute value of the Jacobian**:

$$|\ \mathrm{J}\ | \quad = \left|\ determinant \begin{bmatrix} \dfrac{\partial h_1^{-1}(u_1, u_2)}{\partial u_1} & \dfrac{\partial h_1^{-1}(u_1, u_2)}{\partial u_2} \\ \dfrac{\partial h_2^{-1}(u_1, u_2)}{\partial u_1} & \dfrac{\partial h_2^{-1}(u_1, u_2)}{\partial u_2} \end{bmatrix} \right|$$

$$= \left|\ det \begin{bmatrix} \dfrac{\partial y_1}{\partial u_1} & \dfrac{\partial y_1}{\partial u_2} \\ \dfrac{\partial y_2}{\partial u_1} & \dfrac{\partial y_2}{\partial u_2} \end{bmatrix} \right|$$

$$= \left| \frac{\partial h_1^{-1}(u_1, u_2)}{\partial u_1} \times \frac{\partial h_2^{-1}(u_1, u_2)}{\partial u_2} - \frac{\partial h_1^{-1}(u_1, u_2)}{\partial u_2} \times \frac{\partial h_2^{-1}(u_1, u_2)}{\partial u_1} \right|$$

$$= \left| \frac{\partial y_1}{\partial u_1} \times \frac{\partial y_2}{\partial u_2} - \frac{\partial y_1}{\partial u_2} \times \frac{\partial y_2}{\partial u_1} \right| \qquad \neq\ 0$$

then the joint pdf of U_1 and U_2 is

$$f_{U_1,U_2}(u_1, u_2) = f_{Y_1,Y_2}(\ h_1^{-1}(u_1, u_2), \quad h_2^{-1}(u_1, u_2) \ ) \times |J|$$

Example 1: Let Y_1 and Y_2 be independent random variables with $Y_1 \sim N(0, 1)$ and $Y_2 \sim N(0, 1)$.

Let $U_1 = \dfrac{Y_1 + Y_2}{2}$ and $U_2 = \dfrac{Y_2 - Y_1}{2}$

(a) Find joint pdf of U_1 and U_2.

(b) Find the marginal pdf of U_1

Sections 6.5-6.6 of the textbook

Example 2: Let Y_1 and Y_2 be independent **exponential** random variables, both with **mean** $\beta = 1$.

Let $U_1 = \frac{Y_1}{Y_2}$. Find the pdf of U_1.

Sections 6.5-6.6 of the textbook

Section 6.5: The Method of Moment-Generating Functions

Recall:

A **special expected value** that is quite useful is the moment-generating function (**mgf**).

Definitions 3.14 and 4.14: (Moment-Generating Function)

Let Y be a random variable. The moment-generating function (mgf) of Y is $M(t)$, where

$$M_Y(t) = E(e^{tY}) = \begin{cases} \sum_{\text{all } y} [\, e^{ty}\, p(y) \,] & \text{if } Y \text{ is a } \textbf{discrete} \\ \int_{-\infty}^{\infty} [\, e^{ty}\, f(y) \,] dy & \text{if } Y \text{ is a } \textbf{continuous} \end{cases}$$

for $-b \le t \le b$ where $b(> 0)$ is constant.

Recall:

Distribution of Y with Parameter(s)	**mgf of Y, $M_Y(t)$**	**pmf or pdf of Y**
Binomial (n, p)	$(pe^t + 1 - p)^n$	$\binom{n}{y} p^y (1-p)^{n-y}$
Geometric (p)	$\dfrac{pe^t}{1-(1-p)e^t}$	$p(1-p)^{y-1}$
Negative Binomial (r, p)	$\left[\dfrac{pe^t}{1-(1-p)e^t}\right]^r$	$\binom{y-1}{r-1} p^r (1-p)^{y-r}$
Hypergeometric (N, r, n)	Does not exist	$\dfrac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$
Poisson (λ)	$e^{\lambda(e^t - 1)}$	$\dfrac{e^{-\lambda}\lambda^y}{y!}$
Continuous Uniform (θ_1, θ_2)	$\dfrac{e^{t\theta_2} - e^{t\theta_1}}{t(\theta_2 - \theta_1)}$	$\dfrac{1}{\theta_2 - \theta_1}$
Normal (μ, σ^2)	$e^{\left[\mu t + \frac{1}{2}\sigma^2 t^2\right]}$	$\dfrac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)}$
Gamma (α, β)	$(1-\beta t)^{-\alpha}$	$\dfrac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1} e^{-y/\beta}$
Exponential (β)	$(1-\beta t)^{-1}$	$\dfrac{1}{\beta} e^{-y/\beta}$
Chi-square (v)	$(1-2t)^{-v/2}$	$\dfrac{1}{\Gamma(v/2)\, 2^{v/2}} y^{v/2 - 1} e^{-y/2}$

We can extend this to the multivariate case. The mgf can be used to recognize the probability distribution of function of random variables. However, this method is **limited** to known probability distributions.

Theorem 6.1:

Let $M_X(t)$ and $M_Y(t)$ are moment-generating functions of X and Y, respectively, random variables. **If $M_X(t) = M_Y(t)$ for all the values of t, then X and Y have the same probability distribution.**

Example 3: Let Y be the number of successes in a binomial experiment with n independent trials and p probability of success.

Let $U = n - Y$.

Find the probability distribution of U using mgf.

Example 4: The length of time necessary to tune up a car is exponentially distributed with a mean of 0.5 hour. If two cars are waiting for a tune up and the service times are independent,
(a) find the probability distribution of total time for the two tune ups.
(b) find the probability that total time for the two tune ups will exceed 1.5 hours.

Theorem 6.2: Mgf of sum of independent random variables
Let $Y_1, Y_2, \dots, Y_n$ be **independent** random variables with mgf $M_{Y_1}(t), M_{Y_2}(t), \dots, M_{Y_n}(t)$, respectively.
If $U = Y_1 + Y_2 + \dots + Y_n$,
then mgf of U is

$$M_U(t) = M_{Y_1}(t) \times M_{Y_2}(t) \times \dots \times M_{Y_n}(t)$$

because,

Example 5:

A shipping company handles containers in three different sizes:

(1) 27 ft^3, (2) 125 ft^3, and (3) 512 ft^3.

Let Y_i $(i = 1, 2, 3)$ denote the number of type i containers shipped during a given week and each has the following distributions:

$Y_1 \sim Poisson\ (\lambda_1 = 200)$ $\Longrightarrow$ **mgf of Y_1 is, $M_{Y_1}(t) =$**

$Y_2 \sim Poisson\ (\lambda_2 = 250)$ $\Longrightarrow$ **mgf of Y_2 is, $M_{Y_2}(t) =$**

$Y_3 \sim Poisson\ (\lambda_3 = 100)$ $\Longrightarrow$ **mgf of Y_3 is, $M_{Y_3}(t) =$**

The number of each type of containers shipped is independent from others.

Find the probability distribution of total number of containers shipped per week.

❖ If $Y_i \sim$ **Poisson**(λ_i) for $i = 1, 2, \ldots, n$ and all Y_i **are independent,** then

Theorem 6.3: pdf of <u>linear function of independent Normal</u> random variables

Let $Y_1, Y_2, \ldots, Y_n$ be **independent** random variables each with a **normal** distribution, $E(Y_i) = \mu_i$ and $Var(Y_i) = \sigma_i^2$, for $i = 1,2, \ldots n$.

Let $a_1, a_2, \ldots, a_n$ be constants.

$$\text{If } U = \sum_{i=1}^{n} a_i Y_i = a_1 Y_1 + a_2 Y_2 + \ldots + a_n Y_n \text{, then}$$

U has a **Normal** distribution with **mean** $= \mathrm{E(U)} = \sum_{i=1}^{n} a_i \mu_i$ and **variance** $= \mathrm{Var(U)} = \sum_{i=1}^{n} a_i^2 \sigma_i^2$

Example 6:

Five automobiles of the same type are to be driven on a 300-mile trip. The first two will use an economy brand of gasoline, and the other three will use a name brand.
Let Y_1, Y_2, Y_3, Y_4, and Y_5 be the observed fuel efficiencies (mpg) for the five cars.
Suppose these variables are independent and normally distributed with

$$\mu_1 = \mu_2 = 20, \quad \mu_3 = \mu_4 = \mu_5 = 21, \quad \sigma_1 = \sigma_2 = 2 \text{ and } \sigma_3 = \sigma_4 = \sigma_5 = 1.2$$

U is a measure of the difference in efficiency between economy gas and name-brand gas and defined as

$$U = \frac{Y_1 + Y_2}{2} - \frac{Y_3 + Y_4 + Y_5}{3}$$

Find the pdf of U and then compute $P(U \leq 0)$ and $P(-1 \leq U \leq 1)$.

Sections 6.5-6.6 of the textbook