Multivariate transformations

Transformations via MGF

Idea: If the mgf of a transformed random variable $Y = u(X_1, \ldots, X_n)$ can be identified, then we know the distribution of Y.

Recall if X and Y have mgfs where $M_X(t) = M_Y(t)$ with $t \in (-h, h)$, for some $h > 0$, then X and Y have the same distribution.

General Case:

- Suppose $Y = u(X_1, \ldots, X_n)$ where $(X_1, \ldots, X_n) \sim f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$

- Then mgf of random variable Y is $M_Y(t) = \mathrm{E}e^{tY} = \mathrm{E}e^{tu(X_1,\ldots,X_n)}$, where

$$M_Y(t) = \begin{cases} \mathrm{E}e^{tu(X_1,\ldots,X_n)} = \int \cdots \int e^{tu(x_1,\ldots,x_n)} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) dx_1, \ldots, dx_n & \text{(continuous case)} \\ \mathrm{E}e^{tu(X_1,\ldots,X_n)} = \sum_{(x_1,\ldots,x_n)} e^{tu(x_1,\ldots,x_n)} f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) & \text{(discrete case)} \end{cases}$$

- May recognize the mgf and hence the distribution of Y (if the mgf can be tractably computed)

Most Important Case: sums of independent r.v.s

- Suppose $Y = X_1 + X_2$ with independent X_1, X_2

- Recall that the mgf of random variable Y is then

$$M_Y(t) = \mathrm{E}e^{tY} = \mathrm{E}e^{t(X_1+X_2)} = \mathrm{E}e^{tX_1}e^{tX_2} = \mathrm{E}e^{tX_1}\mathrm{E}e^{tX_2} = M_{X_1}(t)M_{X_2}(t),$$

  assuming the mgfs exist. Note the mgf's of X_1 and X_2 are evaluated at the same t because there is but a single random variable Y

- May again recognize the mgf and hence the distribution of Y. This happens quite often with independent sums.

Multivariate transformations

Transformations via MGF: examples

Suppose X and Y are independent. Let $S = X + Y$ and find the mgf of S when

- $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ (both normally distributed)

$$M_S(t) \stackrel{\text{def}}{=} \mathbb{E}\left[e^{tS}\right] = \mathbb{E}\left[e^{t(X+Y)}\right] \underset{\downarrow}{=} \mathbb{E}\left[e^{tX}\right] \mathbb{E}\left[e^{tY}\right]$$

X and Y are independent

$$= e^{\mu_1 t + \frac{\sigma_1^2 t^2}{2}} \, e^{\mu_2 t + \frac{\sigma_2^2 t^2}{2}}$$

$$= e^{(\mu_1+\mu_2)t + [\sigma_1^2+\sigma_2^2] t^2/2}$$

$$\Longrightarrow \boxed{S \sim \text{Normal}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)}$$

Question: What If $S^* = X - Y$?

$$\Longrightarrow \boxed{S^* \sim \text{Normal}(\mu_1-\mu_2, \sigma_1^2+\sigma_2^2)}$$

- $X \sim$Poisson(λ_1) and $Y \sim$Poisson(λ_2)

$\Longrightarrow$ Find the dist. of $S = X+Y$.

$$M_S(t) = \mathbb{E}\left[e^{t(X+Y)}\right] \underset{\downarrow}{=} \mathbb{E}\left[e^{tX}\right] \mathbb{E}\left[e^{tY}\right]$$

X and Y are independent

$$= \underbrace{e^{\lambda_1(e^t-1)} e^{\lambda_2(e^t-1)}}$$

$$\Longrightarrow S \sim \text{Poisson}(\lambda_1+\lambda_2) = e^{(\lambda_1+\lambda_2)(e^t-1)}$$

$X \sim$ Poisson(λ)

$$M_X(t) = e^{\lambda(e^t-1)}$$

$$M_X(t) = \mathbb{E}\left[e^{tX}\right] = \sum_{x=0}^{\infty} \frac{e^{tx} e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \underbrace{\sum_{x=0}^{\infty} \frac{[\lambda e^t]^x}{x!}} = e^{-\lambda} e^{\lambda e^t} = \boxed{e^{\lambda(e^t-1)}}$$

- X, Y are both Exponential(β) distributed

$\text{Exp}(\beta) \equiv \text{Gamma}(1, \frac{1}{\beta})$

$\Longrightarrow S = X+Y$. Find the dist. of $S = ?$

$$M_S(t) \underset{\downarrow}{=} \underbrace{\mathbb{E}\left[e^{tX}\right] \mathbb{E}\left[e^{tY}\right]} = \left(\frac{\beta}{\beta - t}\right)\left(\frac{\beta}{\beta - t}\right)$$

X and Y are independent

$$= \frac{\beta^2/\beta^2}{(\beta-t)^2/\beta^2}$$

$$= \frac{1}{(1 - \frac{1}{\beta}t)^2}$$

$$\frac{(\beta-t)^2}{\beta^2} = \left(\frac{\beta-t}{\beta}\right)^2$$

$$\Longrightarrow S \sim \text{Gamma}(2, \tfrac{1}{\beta})$$

Recall: If $W \sim \text{gamma}(\alpha,\beta)$

$$M_W(t) = \frac{1}{(1-\beta t)^{\alpha}}$$

$X \sim Exp(\beta) \Rightarrow f_X(x) = \beta e^{-\beta x} = \frac{1}{\frac{1}{\beta}\Gamma(1)} x^{1-1} e^{-x/\frac{1}{\beta}} = f_Y(y)$

$Y \sim Gamma(1, \frac{1}{\beta})$

Multivariate transformations

Transformations via MGF: examples

The same technique shows that: for *independent* $X_1, \ldots, X_n$,

1. if each $X_i \sim N(\mu_i, \sigma_i^2)$, then

$$S = \sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

2. if each $X_i \sim$Binominal(n_i, p) (common p), then

$$S = \sum_{i=1}^{n} X_i \sim \text{ Binomial}\left(\sum_{i=1}^{n} n_i, p\right)$$

3. if each $X_i \sim$Gamma(α_i, β) (common β), then

Gamma(α, β) $f_X(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$

$$S = \sum_{i=1}^{n} X_i \sim \text{ Gamma}\left(\sum_{i=1}^{n} \alpha_i, \beta\right)$$

$X_i \sim Exp(\beta)$ $f_X(x) = \beta e^{-\beta x}$

$X_i \sim Exp(\beta) \sim Gamma(1, \frac{1}{\beta})$

   (a) if each $X_i \sim$Exponential$(\beta) \sim$Gamma$(1, \beta)$, then $S = \sum_{i=1}^n X_i \sim$Gamma$(n, \beta)$ $\frac{1}{\beta}$ $\frac{1}{\beta}$

   (b) if each $X_i \sim \chi^2_{\nu_i} \sim$Gamma$(\nu_i/2, 2)$, then $S = \sum_{i=1}^n X_i \sim \chi^2_{\sum_{i=1}^n \nu_i}$

$X \sim \chi^2_{\nu} \Rightarrow M_X(t) = (1-2t)^{-\nu/2}$, $Y \sim Gamma(\nu/2, 2) \Rightarrow M_Y(t) = (1-2t)^{-\nu/2}$

4. if each $X_i \sim$Possion(λ_i), then

$$S = \sum_{i=1}^{n} X_i \sim \text{ Poisson}\left(\sum_{i=1}^{n} \lambda_i\right)$$

5. if each $X_i \sim$Neg-Binomial(r_i, p) (common p), then

$$S = \sum_{i=1}^{n} X_i \sim \text{ Neg-Binomial}\left(\sum_{i=1}^{n} r_i, p\right)$$

   (a) if each $X_i \sim$Geometric$(p) \sim$Neg-Binomial$(1, p)$, then $S = \sum_{i=1}^n X_i \sim$Neg-Binomial$(n, p)$

Multivariate transformations

Multivariate continuous case

The final technique for determining distributions of transformed random variables:

if we transform continuous random variables, using a one-to-one continuously differentiable transformation, we can *directly* find the pdf of the new random variables

Set-up

- Suppose continuous $(X_1, \ldots, X_n)$ has joint pdf $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ with support $\mathcal{A} = \{(x_1, \ldots, x_n) : f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) > 0\}$

- Transformation:

$$\boldsymbol{Y} = \boldsymbol{u}(\boldsymbol{X}) \quad \text{or} \quad \begin{cases} Y_1 = \underbrace{u_1(X_1, \ldots, X_n)}_{\text{new r.v.}} \\ Y_2 = u_2(X_1, \ldots, X_n) \\ \vdots \quad \vdots \quad \vdots \\ Y_n = u_n(X_1, \ldots, X_n) \end{cases}$$

  with $\mathcal{B}$ =support of $f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n)$

$$\{(y_1,\ldots,y_n) : f_{Y_1,\ldots,Y_n}(y_1,\ldots,y_n) > 0\}$$

- Assume transformation is one-to-one with inverse functions

$$\longrightarrow x_i = u_i^{-1}(y_1, \ldots, y_n), \quad i = 1, \ldots, n$$

$n=2$

$$f_{U,V} = f_{X,Y}(h(u,v), h^*(u,v)) \times \underbrace{(\quad)}_{\text{term}}$$

$$f_{Y_1,\ldots,Y_n} = f_{X_1,\ldots,X_n}(u_1^{-1}(y_1,\ldots,y_n),\ldots,u_n^{-1}(y_1,\ldots,y_n))$$

$x_1 = u_1^{-1}(y_1, \ldots, y_n)$
$\vdots$
$x_n = u_n^{-1}(y_1, \ldots, y_n)$

- Define the Jacobian

$$J = \det \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{pmatrix} = \det \begin{pmatrix} \frac{\partial u_1^{-1}(y_1,\ldots,y_n)}{\partial y_1} & \frac{\partial u_1^{-1}(y_1,\ldots,y_n)}{\partial y_2} & \cdots & \frac{\partial u_1^{-1}(y_1,\ldots,y_n)}{\partial y_n} \\ \frac{\partial u_2^{-1}(y_1,\ldots,y_n)}{\partial y_1} & \frac{\partial u_2^{-1}(y_1,\ldots,y_n)}{\partial y_2} & \cdots & \frac{\partial u_2^{-1}(y_1,\ldots,y_n)}{\partial y_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial u_n^{-1}(y_1,\ldots,y_n)}{\partial y_1} & \frac{\partial u_n^{-1}(y_1,\ldots,y_n)}{\partial y_2} & \cdots & \frac{\partial u_n^{-1}(y_1,\ldots,y_n)}{\partial y_n} \end{pmatrix}$$

- If J is continuous and $J \neq 0$ over $\mathcal{B}$ (except possibly on a set with probability zero),

$$f_{Y_1,\ldots,Y_n}(y_1, \ldots, y_n) = f_{X_1,\ldots,X_n}(\boldsymbol{u}^{-1}(y_1, \ldots, y_n))|J|, \qquad (y_1, \ldots, y_n) \in \mathcal{B}$$

- Often only interested in one transformation $Y_1 = u_1(X_1, \ldots, X_n)$

  Then choose convenient definitions to fill out the transformation

  e.g. $Y_2 = X_2, \ldots, Y_n = X_n$

- If transformation is not one-to-one, then we partition $\mathcal{A}$ (the support of $(X_1, \ldots, X_n)$) into sets $\mathcal{A}_i$ where a transformation $\boldsymbol{Y} = \boldsymbol{u}_j(\boldsymbol{X})$ is one-to-one and then add pieces

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \sum_{i=1}^{k} f_{\boldsymbol{X}}(\boldsymbol{u}_i^{-1}(\boldsymbol{y}))|J_i|$$