4 Transformation of a RV

Consider a RV X with PDF $p(X)$.
Consider a transformed variable $Y = g(X)$, where $g(\cdot)$ is an **increasing** function (we consider only the special case of monotonic functions).

– What is the PDF $p(Y)$?

– Consider probability mass of X in the interval (a, b) getting mapped to the probability mass of Y in the interval $(g(a), g(b))$
– Because we assumed increasing $g(\cdot)$, mass conservation holds, i.e., $P(g(a) < Y < g(b)) = P(a < X < b)$

– Consider $q(y)$ as the PDF of Y
– Now, $P(g(a) < Y < g(b)) = \int_{g(a)}^{g(b)} q(y)dy$

– $P(a < X < b) := \int_a^b p(x)dx$
– Substitute $y = g(x)$ in the integral and write the integral in terms of y.
Then, $x = g^{-1}(y)$
$dx = \left(\frac{d}{dy} g^{-1}(y)\right) dy$
– Then, $P(a < X < b) = \int_{g(a)}^{g(b)} p(g^{-1}(y)) \left(\frac{d}{dy} g^{-1}(y)\right) dy$
– This mass conservation holds for *every interval* (a, b), however small it may be.
– Thus, $q(y) = p(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$

– If $g(\cdot)$ is increasing, then (i) $a < b \implies g(a) < g(b)$ and (ii) the derivative $\frac{d}{dy} g^{-1}(y)$ is non-negative. So, the above formula holds good.

– If $g(\cdot)$ is decreasing, then (i) $a < b \implies g(a) > g(b)$ and (ii) the derivative $\frac{d}{dy} g^{-1}(y)$ is negative. In this case, we can negate the derivative and switch the upper and lower limits to retain the same analysis.

– For convenience, $q(y) = p(g^{-1}(y)) \left|\frac{d}{dy} g^{-1}(y)\right|$

• Classic Example 1 : Consider a RV $X \sim U(0, 1)$ (generated by the C/C++ rand() function). Consider the transformation $Y = (-1/\lambda) \log(X)$. What is $q(Y)$?
– Draw a picture
– $y = -(1/\lambda) \log(x) \implies x = \exp(-\lambda y)$. This is the $g^{-1}(\cdot)$ function.
– $\left|\frac{d}{dy} g^{-1}(y)\right| = \lambda \exp(-\lambda y)$
– So, $q(y) = p(g^{-1}(y)) \left|\frac{d}{dy} g^{-1}(y)\right| = \lambda \exp(-\lambda y)$
– Thus, the PDF of Y is the exponential PDF with parameter λ, i.e., mean $= 1/\lambda$

• Classic Example 2: Consider a RV $X \sim U(-a/2, a/2)$. Consider $Y = \exp(X)$. What is $q(Y)$?
– $y = \exp(x) \implies x = \log(y)$. This is the $g^{-1}(\cdot)$ function.
– $\left|\frac{d}{dy} g^{-1}(y)\right| = 1/y$
– So, $q(y) = p(g^{-1}(y)) \left|\frac{d}{dy} g^{-1}(y)\right| = (1/a)(1/y)$
– Thus, the PDF of Y has form $q(y) = 1/(ay)$ for $y \in (\exp(-a/2), \exp(a/2))$

• Classic Example 3 : Consider a RV $X \sim G(0, 1)$ (standard normal distribution). Consider $Y = aX$ with $a > 0$. What is $q(Y)$?

$$y := ax \implies x = y/a \implies g^{-1}(y) = y/a \tag{1}$$
$$\left| \frac{d}{dy} g^{-1}(y) \right| = 1/a \tag{2}$$
$$q(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = p\left(\frac{y}{a}\right) \frac{1}{a} = \frac{1}{a\sqrt{2\pi}} \exp\left(-\frac{y^2}{2a^2}\right) \tag{3}$$

– Thus, $p(Y)$ is also a Gaussian with σ^2 scaled by a factor of a^2

• Classic Example 4 : Consider a RV $X \sim G(0, a^2)$. Consider $Y = b + X$. What is $q(Y)$?

$$y := b + x \implies x = y - b \implies g^{-1}(y) = y - b$$
$$\left| \frac{d}{dy} g^{-1}(y) \right| = 1$$
$$q(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = p(y-b) \cdot 1 = \frac{1}{a\sqrt{2\pi}} \exp\left(-\frac{(y-b)^2}{2a^2}\right)$$

– Thus, $p(Y)$ is also a Gaussian with μ translated by b

• Example 5 : Consider a PDF $P(X)$ as follows:
$P(x) = 0$ for $x \le -1$
$P(x) = 0.5$ for $x \in (-1, 0)$
$P(x) = 0.5 \exp(-x)$ for $x \ge 0$
Consider a transformation function $y = g(x) = x^2$
What is PDF $q(y)$ of Y ?

Transformation function:

$$y := x^2 \implies x = \sqrt{y} \implies g^{-1}(y) = \sqrt{y}$$
$$\left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{2\sqrt{y}}$$

Case 1 : $x \in (-1, 0)$. In this case, $g(\cdot)$ is a *decreasing* function. Mass conversation applies.

$$\text{For} y \in (0,1) : q_1(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = (0.5) \frac{1}{2\sqrt{y}} = \frac{1}{4\sqrt{y}}$$

Case 2 : $x \ge 0$. In this case, $g(\cdot)$ is a *increasing* function. Mass conversation applies.

$$\text{For} y \ge 0 : q_2(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = (0.5 \exp(-\sqrt{y})) \frac{1}{2\sqrt{y}} = \frac{\exp(-\sqrt{y})}{4\sqrt{y}}$$

Desired PDF $q(y) = q_1(y) + q_2(y)$

In the region $y \in (0, 1)$, the probability mass comes from Case 1 as well as Case 2.

Thus, (i) for $y \in (0, 1)$, PDF $q(y) = \frac{1}{4\sqrt{y}} (1 + \exp(-\sqrt{y}))$
(ii) for $y \ge 0$, PDF $q(y) = \frac{\exp(-\sqrt{y})}{4\sqrt{y}}$
Note the step discontinuity at $y = 1$, where the left limit $= \frac{1+\exp(-1)}{4}$ and the right limit $= \frac{\exp(-1)}{4}$

• Classic Example 6 : Let $X \sim G(0, 1)$. Let $Y := X^2$. Then, what is $P(Y)$, defined as the chi-square PDF ?

Transformation function:

$$y := x^2 \implies x = \sqrt{y} \implies g^{-1}(y) = \sqrt{y}$$
$$\left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{2\sqrt{y}}$$

Case 1 : $x \le 0$. In this case, $g(\cdot)$ is a *decreasing* function. Mass conversation applies.

$$\text{For} y \geq 0 : q_1(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = \frac{\exp(-0.5(\sqrt{y})^2)}{\sqrt{2\pi}} \frac{1}{2\sqrt{y}} = \frac{\exp(-0.5y)}{2\sqrt{y2\pi}}$$

Case 2 : $x > 0$. In this case, $g(\cdot)$ is a *increasing* function. Mass conversation applies.

$$\text{For} y > 0 : q_2(y) := \frac{\exp(-0.5y)}{2\sqrt{y2\pi}}$$

Desired the chi-square PDF is $q(y) = q_1(y) + q_2(y) = (1/\sqrt{y2\pi}) \exp(-0.5y)$

• Classic Example 7 : Let X have a Gamma PDF $P(x) = \text{Gamma}(x|\alpha, \beta) = (\beta^\alpha / \Gamma(\alpha)) x^{\alpha-1} \exp(-\beta x)$, where $\alpha > 0, \beta > 0, x > 0$.
Consider the transformation $Z = 1/X$
What is the PDF of Z ?

Transformation function:

$$y := 1/x \implies x = 1/y \implies g^{-1}(y) = 1/y$$
$$\left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{y^2} \text{for} y > 0$$

For $x > 0$, $g(\cdot)$ is a *decreasing* function. Mass conversation applies.

$$\text{For} y \geq 0 : q_1(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = (\beta^\alpha / \Gamma(\alpha)) y^{1-\alpha} \exp(-\beta/y) \frac{1}{y^2} = (\beta^\alpha / \Gamma(\alpha)) y^{-\alpha-1} \exp(-\beta/y)$$

5 More transformations on RVs: Addition, Min, Max

– Given: RV X, Y are *independent* and have PDsF $A(X), B(Y)$.
– Let $Z := X + Y$
– What is the PDF $C(Z)$?

Case 1: Discrete RVs
$C(Z = z) = P(\{(x, y)|X = x, Y = z - x\}) = \sum_x P(X = x, Y = z - x) = \sum_x A(X = x)B(Y = z - x)$ (because of independence)

This operation is called convolution in the discrete case.

Case 2: Continuous RVs
$C(Z \in (z_1, z_2)) = P(\{(x, y)|X = x, Y \in (z_1 - x, z_2 - x)\}) = \int_x P(X = x, Y \in (z_1 - x, z_2 - x))dx$
$= \int_{z_1}^{z_2} \int_x P(X = x, Y = z - x)dxdz$
Thus, $P(Z = z) = \int_x P(X = x, Y = z - x)dx = \int_x A(X = x)B(Y = z - x)dx$ (because of independence)

This operation is called convolution in the continuous case.

– Example: Discrete case.
Integer RV X such that $P(x = 5) = 1$; $P(x \neq 5) = 0$ (One delta function)
Integer RV Y such that $P(y) = 1/7$ for $y = -3, -2, \cdots, 2, 3$; $P(y) = 0$ otherwise
Let $Z := X + Y$
What is $P(Z)$?

– Example: Discrete case.
Integer RV X such that $P(x = 5) = 0.5$; $P(x = 10) = 0.5$; $P(x) = 0$ otherwise (Two delta functions)
Integer RV Y such that $P(y) = 1/7$ for $y = -3, -2, \cdots, 2, 3$; $P(y) = 0$ otherwise
Let $Z := X + Y$
What is $P(Z)$?

– Convolution spreads the probability mass

– Convolution and the central limit theorem

For i.i.d. RVs X_i with an arbitrary PDF $P(X)$, $\lim_{n\to\infty} \sum_i X_i$ has a Gaussian PDF
This implies that the PDF / function obtained by successive convolution $P(x) * P(x) * \cdots$ tends to a Gaussian PDF / function, for any arbitrary $P(X)$

– Addition of $Z_1 \sim G(0, \sigma_1^2)$ and $Z_2 \sim G(0, \sigma_2^2)$. If $Z_3 := Z_1 + Z_2$, then what is $P(Z_3)$?

– Let X_i be *independent* RVs with CDFs $F_i(X_i)$

– Let $V = \max(X_1, \cdots, X_n)$
– Then, $P(V \le x) = P(X_1 \le x, \cdots, X_n \le x) = P(X_1 \le x) \cdots P(X_n \le x)$ (because of independence)
$= F_1(x) \cdots F_n(x)$
– Thus, the CDF of V is $H(x) = P(V \le x) = F_1(x) \cdots F_n(x)$

– Let $U = \min(X_1, \cdots, X_n)$
– Then, $P(U > x) = P(X_1 > x, \cdots, X_n > x) = P(X_1 > x) \cdots P(X_n > x)$ (because of independence)
$= (1 - F_1(x)) \cdots (1 - F_n(x))$
– Then, the CDF of U is $H(x) = 1 - P(U > x) = 1 - (1 - F_1(x)) \cdots (1 - F_n(x))$