Outline

Transformations

Distribution Function Technique

Change of Variables

Univariate Gaussian Distributions
Outline

Transformations

Distribution Function Technique

Change of Variables

Univariate Gaussian Distributions
Review of Distributions

We have seen many named distributions:

- Bernoulli - a coin flip
- Binomial - a series of coin flips
- Gaussian/Normal - height
- Poisson - amount of (e)mail(s) you receive daily
Review of Distributions cont.

For every distribution there are several things to keep in mind:

- Discrete or Continuous
- Parameters
- Probability mass/density function
- Support - nonzero parts
- Expectation or Mean
- Variance
Motivating Example

For named distribution we have a lot of information.

\[ X \sim \mathcal{N}(0, 1) \]

But what about \( X^2 \)? or \( \log(X) \)?

More generally if I have a a function \( U(X) \), what can our information about \( X \) tell us about \( U(X) \)?
Approaches

- **Discrete**
  - Direct Change

- **Continuous**
  - Distribution Function Technique
  - Change of Variables
Suppose $X$ is distributed according to any discrete distribution, and we have an invertible function $U(X) = Y$, then

$$P(Y = y) = P(U(X) = y) = P(X = U^{-1}(y)).$$

Implies we can use $X$’s pmf on the event $U^{-1}(y)$. 
Outline

Transformations

Distribution Function Technique

Change of Variables

Univariate Gaussian Distributions
Continuous Case

For a continuous random variable $X$, a function $Y = U(X)$:

1. Find the cdf:

   $F_Y(y) = P(Y \leq y)$

2. Differentiate the cdf $F_Y(y)$ to get the pdf $f_Y(y)$:

   $f_Y(y) = \frac{d}{dy} F_Y(y)$. 
Example 1 - simple pdf

Let \( X \) be a continuous random variable defined on the interval \([0, 1]\) with pdf

\[
f_X(x) = 3x^2.
\]

What is the pdf of the random variable \( Y = X^2 \)?
Step 1: Find the cdf.

\[ F_Y(y) = P(Y \leq y) \]
\[ = P(X^2 \leq y) \]
\[ = P(X \leq y^{1/2}) \]
\[ = F_X(y^{1/2}) \]
\[ = \int_0^{y^{1/2}} 3t^2 \, dt \]
\[ = \left[ t^3 \right]_{t=0}^{y^{1/2}} \]
\[ F_Y(y) = y^{3/2}, \quad y \in [0, 1] \]
Outline

Transformations

Distribution Function Technique

Change of Variables

Univariate Gaussian Distributions
Change of Variables Steps

$X$ is a univariate random variable (r.v.) with states $x \in [a, b]$ and pdf $f(x)$. Another r.v. $Y = U(X)$, where $U$ is an invertible function. What is pdf $f(y)$?

Steps:

1. Transform cdf of $Y$ into cdf of $X$.
2. Differentiate cdf to get pdf.
1. Transform cdf of $Y$ into cdf of $X$.

By definition of cdf:

$$F_Y(y) = P(Y \leq y) = P(U(X) \leq y)$$

Assume $U$ is strictly increasing, then $U^{-1}$ is also strictly increasing.

$$P(U(X) \leq y) = P(U^{-1}(U(X)) \leq U^{-1}(y))$$

$$= P(X \leq U^{-1}(y))$$
2. Differentiate cdf to get pdf.

Based on definition of the cdf of $X$,

$$F_Y(y) = P(X \leq U^{-1}(y)) = \int_{a}^{U^{-1}(y)} f(x) \, dx$$

Differentiate with respect to $y$,

$$f(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \int_{a}^{U^{-1}(y)} f(x) \, dx$$
\[ \int f(x)dx = \int f(U^{-1}(y))U^{-1'}(y)dy \]

\[ f(y) = \frac{d}{dy} \int_a^{U^{-1}(y)} f_x(U^{-1}(y))U^{-1'}(y)dy \]

\[ = f_x(U^{-1}(y)) \cdot \left( \frac{d}{dy} U^{-1}(y) \right) . \]

For both increasing and decreasing \( U \),

\[ f(y) = f_x(U^{-1}(y)) \cdot \left| \frac{d}{dy} U^{-1}(y) \right| . \]
Example 2: Univariate Normal

**Theorem**

Suppose $X \sim N(\mu, \sigma^2)$ and $Z = U(X) = \frac{X-\mu}{\sigma}$. Then $Z \sim N(0, 1)$.

**Analysis:**

$$f(z) = f_x(U^{-1}(z)) \cdot \left| \frac{d}{dz} U^{-1}(z) \right|$$
Example 2: Univariate Normal Cont.

Proof: \( f_X(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} \),

\[
x = U^{-1}(z) = \sigma z + \mu, \quad \frac{d}{dz} U^{-1}(z) = \sigma.
\]

\[
f(z) = f_X(U^{-1}(z)) \cdot \left| \frac{d}{dz} U^{-1}(z) \right| = f_X(\sigma z + \mu) \cdot |\sigma| \quad = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} z^2} \cdot |\sigma| \quad = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2}
\]
Multivariate Change of Variables

**Theorem**

Let $X$ be a multivariate continuous r.v., $f_x(x)$ be the pdf. If the vector-valued function $y = U(x)$ is differentiable and invertible for all values within the domain of $x$, then for corresponding values of $y$, the pdf of $Y = U(X)$ is given by

$$f(y) = f_x(U^{-1}(y)) \cdot \left| \det \left( \frac{\partial}{\partial y} U^{-1}(y) \right) \right|$$

(Univariate: $f(y) = f_x(U^{-1}(y)) \cdot | \frac{d}{dy} U^{-1}(y) |$)
Example 3: Multivariate Gaussian

Let $A$ be an invertible $p \times p$ matrix, $\mu \in \mathbb{R}^{p \times 1}$, and $Z = (Z_1, \ldots, Z_p)' \in \mathbb{R}^{p \times 1}$ be independent standard normal r.v.’s $\{Z_j\} \sim N(0, 1)$, with joint pdf

$$f_Z(z) = (2\pi)^{-\frac{p}{2}} e^{-\frac{z'z}{2}}.$$  

Then $X = g(Z) = \mu + AZ \sim N(\mu, C)$, where

$$C = E(X - \mu)(X - \mu)'$$

$$= E(AZ)(AZ)'$$

$$= E[AZZ'A'] = AA'$$
Example 3: Multivariate Gaussian Cont.

Proof: \( f(x) = f_z(g^{-1}(x)) \cdot | \det(\frac{\partial}{\partial x} g^{-1}(x)) | \)

\[ g^{-1}(x) = A^{-1}(x - \mu), \quad \frac{\partial}{\partial x} g^{-1}(x) = A^{-1} \]

\[ f(x) = f_z(A^{-1}(x - \mu)) \cdot | \det A |^{-1} \]
\[ = (2\pi)^{-\frac{p}{2}} e^{-\frac{1}{2} (x-\mu)'(A^{-1})'A^{-1}(x-\mu)} / \sqrt{\det AA'} \]
\[ = \frac{1}{\sqrt{(2\pi)^p |\det C|}} e^{-\frac{1}{2} (x-\mu)'C^{-1}(x-\mu)} \]
Outline

Transformations

Distribution Function Technique

Change of Variables

Univariate Gaussian Distributions
Example 4: Chi-Square

Let $X \sim \mathcal{N}(0, 1)$ and $Y = X^2$. The square function is not one-to-one on the whole real line (i.e. it’s inverse only is defined for positive numbers).

However, $X^2 \leq y \iff X \in [-\sqrt{y}, \sqrt{y}]$. Then

$$F_Y(y) = P(X^2 \leq y)$$

$$= P(-\sqrt{y} \leq X \leq \sqrt{y})$$

$$= \Phi(\sqrt{y}) - \Phi(-\sqrt{y}) = 2\Phi(\sqrt{y}) - 1$$
Example 5: Log-Normal

Once again let $X \sim \mathcal{N}(0, 1)$, and $Y = e^X$. Since the exponential function is strictly increasing and is one-to-one on the whole real line, then we can just apply the change of variable formula. Recall that $x = \log y$, and $dy/dx = e^x$. We have that

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right| = \varphi(x) \frac{1}{e^x} = \varphi(\log y) \frac{1}{y}, \quad y > 0$$