Questions?
(Dose response curve and conditional density functions)
Where We’ve Been, Where We’re Going

Finishing Up Yesterday:

3) Independence, Expectation, Covariance
4) Properties of Sums of Random Variables
5) The Multivariate Normal Distribution and You

Today:

1) Properties of Expectations, Changing Coordinates
2) Moment Generating Functions
Definition

*Two random variables $X$ and $Y$ are independent if for any two sets of real numbers $A$ and $B$,*

$$P\{X \in A, Y \in B\} = P\{X \in A\} \cdot P\{Y \in B\}$$

Equivalently we will say $X$ and $Y$ are independent if,

$$f(x, y) = f_X(x) \cdot f_Y(y)$$

If $X$ and $Y$ are not independent, we will say they are dependent.
Definition
Two random variables $X$ and $Y$ are independent if for any two sets of real numbers $A$ and $B$,

$$P\{X \in A, Y \in B\} = P\{X \in A\} P\{Y \in B\}$$
Definition

Two random variables $X$ and $Y$ are independent if for any two sets of real numbers $A$ and $B$,

$$P\{X \in A, Y \in B\} = P\{X \in A\} P\{Y \in B\}$$

Equivalently we will say $X$ and $Y$ are independent if,
Definition

Two random variables $X$ and $Y$ are independent if for any two sets of real numbers $A$ and $B$,

$$P\{X \in A, Y \in B\} = P\{X \in A\} P\{Y \in B\}$$

Equivalently we will say $X$ and $Y$ are independent if,

$$f(x, y) = f_X(x)f_Y(y)$$
Definition

Two random variables $X$ and $Y$ are independent if for any two sets of real numbers $A$ and $B$,

$$P\{X \in A, Y \in B\} = P\{X \in A\} P\{Y \in B\}$$

Equivalently we will say $X$ and $Y$ are independent if,

$$f(x, y) = f_X(x) f_Y(y)$$

If $X$ and $Y$ are not independent, we will say they are dependent
Conditional Distribution

If $X$ and $Y$ are independent, then

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

$$= \frac{f_X(x)f_Y(y)}{f_Y(y)}$$

$$= f_X(x)$$

In words: the distribution of $X$ does not change as levels of $Y$ change.
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = \begin{cases} x + y, & \text{if } x \in [0, 1], y \in [0, 1] \\ 0, & \text{otherwise} \end{cases}$$

$$f_X(x) = \int_0^1 (x + y) \, dy = xy + y^2/2 \bigg|_0^1 = x + 1/2$$

$$f_Y(y) = 1/2 + y$$

Intuition: at different levels of $X$ the distribution on $Y$ behaves differently. $X$ provides information about $Y$. 
A (Simple) Example of Dependence
Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = x + y, \text{ if } x \in [0, 1], y \in [0, 1]$$
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = \begin{cases} x + y, & \text{if } x \in [0, 1], y \in [0, 1] \\ 0, & \text{otherwise} \end{cases}$$

Intuition: at different levels of $X$ the distribution on $Y$ behaves differently.
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = \begin{cases} 
  x + y , & \text{if } x \in [0, 1], y \in [0, 1] \\
  0 , & \text{otherwise}
\end{cases}$$

Intuition: at different levels of $X$ the distribution on $Y$ behaves differently.

$$f_X(x) = \int_0^1 (x + y) \, dy$$
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = x + y, \text{ if } x \in [0, 1], y \in [0, 1]$$

$$= 0, \text{ otherwise}$$

$$f_X(x) = \int_0^1 (x + y) \, dy$$

$$= xy + \frac{y^2}{2} \bigg|_0^1$$
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = x + y, \text{ if } x \in [0, 1], y \in [0, 1]$$
$$= 0, \text{ otherwise}$$

$$f_X(x) = \int_0^1 (x + y) \, dy$$
$$= xy + \frac{y^2}{2} \bigg|_0^1$$
$$= x + \frac{1}{2}$$

Intuition: at different levels of $X$ the distribution on $Y$ behaves differently.
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = \begin{cases} x + y, & \text{if } x \in [0, 1], y \in [0, 1] \\ 0, & \text{otherwise} \end{cases}$$

$$f_X(x) = \int_0^1 (x + y) \, dy$$

$$= xy + \frac{y^2}{2} \bigg|_0^1$$

$$= x + \frac{1}{2}$$

$$f_Y(y) = \frac{1}{2} + y$$
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = \begin{cases} x + y , & \text{if } x \in [0, 1], y \in [0, 1] \\ 0 , & \text{otherwise} \end{cases}$$

Intuition: at different levels of $X$ the distribution on $Y$ behaves differently.

$X$ provides information about $Y$. 

$$f(x, y) = x + y$$
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = x + y, \text{ if } x \in [0,1], y \in [0,1]$$
$$= 0, \text{ otherwise}$$

$$f(x, y) = x + y$$
$$f_X(x)f_Y(y) = \left(\frac{1}{2} + x\right)\left(\frac{1}{2} + y\right)$$

Intuition: at different levels of $X$ the distribution on $Y$ behaves differently.

$X$ provides information about $Y$. 

Justin Grimmer (Stanford University)
Methodology I
September 16th, 2015 6 / 42
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = x + y \text{, if } x \in [0, 1], y \in [0, 1]$$
$$= 0 \text{, otherwise}$$

$$f(x, y) = x + y$$

$$f_X(x)f_Y(y) = \left(\frac{1}{2} + x\right)\left(\frac{1}{2} + y\right)$$
$$= \frac{1}{4} + \frac{x + y}{2} + xy$$
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = x + y, \text{ if } x \in [0, 1], y \in [0, 1]$$

$$= 0, \text{ otherwise}$$

$$f(x, y) = x + y$$

$$f_X(x)f_Y(y) = \left(\frac{1}{2} + x\right)\left(\frac{1}{2} + y\right)$$

$$= \frac{1}{4} + \frac{x + y}{2} + xy$$

Intuition: at different levels of $X$ the distribution on $Y$ behaves differently.
A (Simple) Example of Dependence

Suppose $X$ and $Y$ are jointly continuous and that

$$f(x, y) = x + y, \text{ if } x \in [0, 1], y \in [0, 1]$$
$$= 0, \text{ otherwise}$$

$$f(x, y) = x + y$$
$$f_X(x)f_Y(y) = \left(\frac{1}{2} + x\right)\left(\frac{1}{2} + y\right)$$
$$= \frac{1}{4} + \frac{x + y}{2} + xy$$

Intuition: at different levels of $X$ the distribution on $Y$ behaves differently. $X$ provides information about $Y$
Expectation

Definition

For jointly continuous random variables \( X \) and \( Y \) define,

\[
E[X] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x, y) \, dx \, dy
\]

\[
E[Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x, y) \, dx \, dy
\]

\[
E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) \, dx \, dy
\]

Proposition

Suppose \( g : \mathbb{R}^2 \to \mathbb{R} \) (that isn’t crazy). Then,

\[
E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)f(x, y) \, dx \, dy
\]
Covariance

Definition

*For jointly continuous random variables $X$ and $Y$ define, the covariance of $X$ and $Y$ as,*

\[
\text{cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]
\]

*Define the correlation of $X$ and $Y$ as,*

\[
\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}
\]
Covariance

Definition

For jointly continuous random variables $X$ and $Y$ define, the covariance of $X$ and $Y$ as,

$$\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$$
Covariance

Definition

For jointly continuous random variables $X$ and $Y$ define, the covariance of $X$ and $Y$ as,

$$
$$
Covariance

Definition

For jointly continuous random variables $X$ and $Y$ define, the covariance of $X$ and $Y$ as,

$$\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$$


Covariance

Definition

For jointly continuous random variables $X$ and $Y$ define, the covariance of $X$ and $Y$ as,

\[
\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])]
\]

\[
\]

\[
\]

\[
= E[XY] - E[X]E[Y]
\]
Covariance

Definition

*For jointly continuous random variables* \( X \) *and* \( Y \) *define, the covariance of* \( X \) *and* \( Y \) *as,*

\[
\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])]
\]

\[
\]

\[
\]

\[
= E[XY] - E[X]E[Y]
\]

*Define the correlation of* \( X \) *and* \( Y \) *as,*
Covariance

Definition

For jointly continuous random variables $X$ and $Y$ define, the covariance of $X$ and $Y$ as,

$$
cov(X, Y) = E[(X - E[X])(Y - E[Y])]
= E[XY] - E[X]E[Y]
$$

Define the correlation of $X$ and $Y$ as,

$$
cor(X, Y) = \frac{cov(X, Y)}{\sqrt{Var(X)Var(Y)}}
$$
Some Observations

Variance is the covariance of a random variable with itself.
Some Observations

Variance is the covariance of a random variable with itself.

\[
\text{cov}(X, X) = E[XX] - E[X]E[X] \\
= E[X^2] - E[X]^2
\]
Some Observations

Variance is the covariance of a random variable with itself.

\[ \text{cov}(X, X) = E[XX] - E[X]E[X] = E[X^2] - E[X]^2 \]

Correlation measures the linear relationship between two random variables
Some Observations

Variance is the covariance of a random variable with itself.

\[
\text{cov}(X, X) = E[XX] - E[X]E[X] \\
= E[X^2] - E[X]^2
\]

Correlation measures the linear relationship between two random variables
Suppose \( X = Y \)
Some Observations

Variance is the covariance of a random variable with itself.

\[
\text{cov}(X, X) = E[XX] - E[X]E[X] \\
= E[X^2] - E[X]^2
\]

Correlation measures the linear relationship between two random variables

Suppose \( X = Y \)

\[
\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}
\]
Some Observations

Variance is the covariance of a random variable with itself.

$$\text{cov}(X, X) = E[XX] - E[X]E[X]$$

$$= E[X^2] - E[X]^2$$

Correlation measures the linear relationship between two random variables

Suppose $X = Y$

$$\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

$$= \frac{\text{Var}(X)}{\text{Var}(X)}$$
Some Observations

Variance is the covariance of a random variable with itself.

\[ \text{cov}(X, X) = E[XX] - E[X]E[X] \]
\[ = E[X^2] - E[X]^2 \]

Correlation measures the linear relationship between two random variables

Suppose \( X = Y \)

\[ \text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \]
\[ = \frac{\text{Var}(X)}{\text{Var}(X)} \]
\[ = 1 \]
Some Observations

Variance is the covariance of a random variable with itself.

\[
\text{cov}(X, X) = E[XX] - E[X]E[X] \\
= E[X^2] - E[X]^2
\]

Correlation measures the linear relationship between two random variables

Suppose \( X = Y \)

\[
\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \\
= \frac{\text{Var}(X)}{\text{Var}(X)} \\
= 1
\]

Suppose \( X = -Y \)
Some Observations

Variance is the covariance of a random variable with itself.

\[
\text{cov}(X, X) = E[XX] - E[X]E[X] \\
= E[X^2] - E[X]^2
\]

Correlation measures the linear relationship between two random variables

Suppose \( X = Y \)

\[
\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \\
= \frac{\text{Var}(X)}{\text{Var}(X)} \\
= 1
\]

Suppose \( X = -Y \)

\[
\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}
\]
Some Observations

Variance is the covariance of a random variable with itself.

\[
\text{cov}(X, X) = E[XX] - E[X]E[X]
\]

\[
= E[X^2] - E[X]^2
\]

Correlation measures the linear relationship between two random variables.

Suppose \( X = Y \)

\[
\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}
\]

\[
= \frac{\text{Var}(X)}{\text{Var}(X)}
\]

\[
= 1
\]

Suppose \( X = -Y \)

\[
\text{cor}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}
\]

\[
= \frac{-\text{Var}(X)}{\text{Var}(X)}
\]

\[
= -1
\]
Correlation is Between -1 and 1

|\text{cor}(X, Y)| \leq 1

- Proof 1: Variance trick
- Proof 2: Cauchy-Schwartz Inequality
  - “Inner product” of any two vectors $X$ and $Y$ is less than or equal to the length of vector $X$ times the length of vector $Y
Example: $X + Y$
Suppose $X$ and $Y$ have pdf $x + y$ for $x, y \in [0, 1]$. 

$\text{Cov}(X, Y) = \mathbb{E}[XY] = \int_0^1 \int_0^1 xy(x + y) \, dx \, dy = \int_0^1 \int_0^1 (x^2 y + y^2 x) \, dx \, dy = \int_0^1 (y^3 + y^2) \, dy = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$

$\mathbb{E}[X] = \int_0^1 \int_0^1 x(x + y) \, dx \, dy = \frac{7}{12}$
Example: $X + Y$

Suppose $X$ and $Y$ have pdf $x + y$ for $x, y \in [0, 1]$.

\[
\text{Cov}(X, Y) = E[XY] = \int_0^1 \int_0^1 xy(x + y) \, dx \, dy = \int_0^1 \int_0^1 (x^2 y + y^2 x) \, dx \, dy = \int_0^1 (y^3 + y^2 x) \, dy = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}
\]

\[
E[X] = \int_0^1 \int_0^1 x(x + y) \, dx \, dy = \frac{7}{12}
\]
Example: $X + Y$

Suppose $X$ and $Y$ have pdf $x + y$ for $x, y \in [0, 1]$.

$\text{Cov}(X, Y)$

\[
E[XY] = \int_0^1 \int_0^1 xy(x + y) \, dx \, dy
\]
Example: $X + Y$

Suppose $X$ and $Y$ have pdf $x + y$ for $x, y \in [0, 1]$.

$\text{Cov}(X, Y)$

\[
E[XY] = \int_0^1 \int_0^1 xy(x + y) \, dx \, dy
\]
\[
= \int_0^1 \int_0^1 (x^2 y + y^2 x) \, dx \, dy
\]
Example: $X + Y$

Suppose $X$ and $Y$ have pdf $x + y$ for $x, y \in [0, 1]$.

$\text{Cov}(X, Y)$

$$E[XY] = \int_0^1 \int_0^1 xy(x + y) \, dx \, dy$$

$$= \int_0^1 \int_0^1 (x^2 y + y^2 x) \, dx \, dy$$

$$= \int_0^1 \left( \frac{y}{3} + \frac{y^2}{2} \right) \, dy$$
Example: $X + Y$

Suppose $X$ and $Y$ have pdf $x + y$ for $x, y \in [0, 1]$. Cov($X, Y$)

\[
E[XY] = \int_0^1 \int_0^1 xy(x + y) \, dx \, dy
\]

\[
= \int_0^1 \int_0^1 (x^2 y + y^2 x) \, dx \, dy
\]

\[
= \int_0^1 \left( \frac{y}{3} + \frac{y^2}{2} \right) \, dy
\]

\[
= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}
\]
Example: $X + Y$

Suppose $X$ and $Y$ have pdf $x + y$ for $x, y \in [0, 1]$.

Cov($X, Y$)

\[
E[XY] = \int_0^1 \int_0^1 xy(x + y) \, dx \, dy
\]
\[
= \int_0^1 \int_0^1 (x^2 y + y^2 x) \, dx \, dy
\]
\[
= \int_0^1 \left( \frac{y}{3} + \frac{y^2}{2} \right) \, dy
\]
\[
= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}
\]

\[
E[X] = \int_0^1 \int_0^1 x(x + y) \, dx \, dy
\]
Example: $X + Y$

Suppose $X$ and $Y$ have pdf $x + y$ for $x, y \in [0, 1]$.

$\text{Cov}(X, Y)$

$$E[XY] = \int_0^1 \int_0^1 xy(x + y) \, dx \, dy$$

$$= \int_0^1 \int_0^1 (x^2 y + y^2 x) \, dx \, dy$$

$$= \int_0^1 \left( \frac{y}{3} + \frac{y^2}{2} \right) \, dy$$

$$= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$$E[X] = \int_0^1 \int_0^1 x(x + y) \, dx \, dy$$

$$= \frac{7}{12}$$
Example: $X + Y$

Suppose $X$ and $Y$ have pdf $x + y$ for $x, y \in [0, 1]$.

$\text{Cov}(X, Y)$

$$E[XY] = \int_0^1 \int_0^1 xy(x + y) \, dx \, dy$$

$$= \int_0^1 \int_0^1 (x^2 y + y^2 x) \, dx \, dy$$

$$= \int_0^1 \left( \frac{y}{3} + \frac{y^2}{2} \right) dy$$

$$= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$$E[Y] = \int_0^1 \int_0^1 y(x + y) \, dx \, dy$$

$$= \int_0^1 \int_0^1 y \, dx \, dy$$

$$= \frac{1}{6}$$
Example: $X + Y$

Suppose $X$ and $Y$ have pdf $x + y$ for $x, y \in [0, 1]$.

\[
\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]
\]

\[
\mathbb{E}[XY] = \int_0^1 \int_0^1 xy(x + y) \, dx \, dy
\]

\[
= \int_0^1 \int_0^1 (x^2y + y^2x) \, dx \, dy
\]

\[
= \int_0^1 \left( \frac{y}{3} + \frac{y^2}{2} \right) dy
\]

\[
= \frac{1}{6} + \frac{1}{6} = \frac{1}{3}
\]

\[
\mathbb{E}[Y] = \int_0^1 \int_0^1 y(x + y) \, dx \, dy
\]

\[
= \frac{7}{12}
\]
Example: $X + Y$

\[
\text{Cov}(X, Y) = E[XY] - E[X]E[Y]
\]
Example: \( X + Y \)

\[
\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}
\]
Example: $X + Y$

\[
\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}
\]

\[
\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}
\]
Example: $X + Y$

\[
\text{Cov}(X, Y) = E[XY] - E[X]E[Y]
\]

\[
= \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}
\]

\[
\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}
\]

\[
= -\frac{1}{144} \cdot \frac{11}{144}
\]
Example: $X + Y$

\[
\text{Cov}(X, Y) = E[XY] - E[X]E[Y]
\]
\[
= \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}
\]

\[
\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}
\]
\[
= \frac{-\frac{1}{144}}{\frac{11}{144}} = -\frac{1}{11}
\]
Sums of Random Variables

Suppose we have a sequence of random variables $X_i$, $i = 1, 2, \ldots, N$. Suppose that they have joint pdf,

$$f(x) = f(x_1, x_2, \ldots, x_n)$$

1) $E[\sum_{i=1}^N X_i] = \sum_{i=1}^N E[X_i]$
2) $\text{var}(\sum_{i=1}^N X_i) = \sum_{i=1}^N \text{var}(X_i) + 2 \sum_{i<j} \text{cov}(X_i, X_j)$
Proposition

Suppose we have a sequence of random variables $X_i$, $i = 1, 2, \ldots, N$. Suppose that they have joint pdf,

$$f(x) = f(x_1, x_2, \ldots, x_n)$$

Then

$$E\left[ \sum_{i=1}^{N} X_i \right] = \sum_{i=1}^{N} E[X_i]$$
Proof.
Proof.

\[ E[\sum_{i=1}^{N} X_i] = E[X_1 + X_2 + \ldots + X_N] \]
Proof.

\[ E[\sum_{i=1}^{N} X_i] = E[X_1 + X_2 + \ldots + X_N] \]

\[ = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1 + x_2 + \ldots + x_N) f(x_1, x_2, \ldots, x_N) dx_1 dx_2 \ldots dx_N \]
Proof.

\[
E\left[\sum_{i=1}^{N} X_i\right] = E[X_1 + X_2 + \ldots + X_N]
\]

\[
= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1 + x_2 + \ldots + x_N)f(x_1, x_2, \ldots, x_N)dx_1 dx_2 \ldots dx_N
\]

\[
= \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1)dx_1 + \int_{-\infty}^{\infty} x_2 f_{X_2}(x_2)dx_2 + \ldots + \int_{-\infty}^{\infty} x_N f_{X_N}(x_N)dx_N
\]
Proof.

\[
E \left[ \sum_{i=1}^{N} X_i \right] = E[X_1 + X_2 + \ldots + X_N]
\]

\[
= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} (x_1 + x_2 + \ldots + x_N) f(x_1, x_2, \ldots, x_N) dx_1 dx_2 \ldots dx_N
\]

\[
= \int_{-\infty}^{\infty} x_1 f_{X_1}(x_1) dx_1 + \int_{-\infty}^{\infty} x_2 f_{X_2}(x_2) dx_2 + \ldots + \int_{-\infty}^{\infty} x_N f_{X_N}(x_N) dx_N
\]

\[
= E[X_1] + E[X_2] + \ldots + E[X_N]
\]
Sums of Random Variable

Proposition

Suppose $X_i$ is a sequence of random variables. Then

\[ \text{var} \left( \sum_{i=1}^{N} X_i \right) = \sum_{i=1}^{N} \text{var}(X_i) + 2 \sum_{i<j} \text{cov}(X_i, X_j) \]
Sums of Random Variable

Proof.
Consider two random variables, $X_1$ and $X_2$. Then,

$$\text{var}(X_1 + X_2) = E[(X_1 + X_2)^2] - (E[X_1] + E[X_2])^2$$

$$= E[X_2^2] + 2E[X_1X_2] - (E[X_1])^2 - 2E[X_1]E[X_2] - E[X_2]^2$$

$$= \text{var}(X_1) + \text{var}(X_2) + 2 \text{cov}(X_1, X_2)$$
Sums of Random Variable

Proof.
Consider two random variables, $X_1$ and $X_2$. Then,

$$\text{var}(X_1 + X_2) = E[(X_1 + X_2)^2] - (E[X_1] + E[X_2])^2$$
Sums of Random Variable

Proof.
Consider two random variables, $X_1$ and $X_2$. Then,

$$\text{var}(X_1 + X_2) = E[(X_1 + X_2)^2] - (E[X_1] + E[X_2])^2$$
$$= E[X_1^2] + 2E[X_1X_2] + E[X_2^2] - (E[X_1])^2 - 2E[X_1]E[X_2] - E[X_2]^2$$

$\text{cov}(X_1, X_2) = \text{var}(X_1) + \text{var}(X_2) + 2\text{cov}(X_1, X_2)$
Sums of Random Variable

Proof.
Consider two random variables, $X_1$ and $X_2$. Then,

$$\text{var}(X_1 + X_2) = \mathbb{E}[(X_1 + X_2)^2] - (\mathbb{E}[X_1] + \mathbb{E}[X_2])^2$$
$$= \mathbb{E}[X_1^2] + 2\mathbb{E}[X_1 X_2] + \mathbb{E}[X_2^2] - (\mathbb{E}[X_1])^2 - 2\mathbb{E}[X_1]\mathbb{E}[X_2] - \mathbb{E}[X_2]^2$$
$$= \underbrace{\mathbb{E}[X_1^2] - (\mathbb{E}[X_1])^2}_{\text{var}(X_1)} + \underbrace{\mathbb{E}[X_2^2] - \mathbb{E}[X_2]^2}_{\text{var}(X_2)} + 2\underbrace{(\mathbb{E}[X_1 X_2] - \mathbb{E}[X_1] \mathbb{E}[X_2])}_{\text{cov}(X_1, X_2)}$$

$$= \text{var}(X_1) + \text{var}(X_2) + 2\text{cov}(X_1, X_2)$$
Sums of Random Variables

Proof.
Consider two random variables, $X_1$ and $X_2$. Then,

\[
\text{var}(X_1 + X_2) = E[(X_1 + X_2)^2] - (E[X_1] + E[X_2])^2
\]
\[
= E[X_1^2] + 2E[X_1X_2] + E[X_2^2] - (E[X_1])^2 - 2E[X_1]E[X_2] - E[X_2]^2
\]
\[
= \underbrace{E[X_1^2] - (E[X_1])^2}_{\text{var}(X_1)} + \underbrace{E[X_2^2] - E[X_2]^2}_{\text{var}(X_2)}
\]
\[
+ 2 \left( E[X_1X_2] - E[X_1]E[X_2] \right) \underbrace{\text{cov}(X_1,X_2)}_{\text{cov}(X_1,X_2)}
\]
\[
= \text{var}(X_1) + \text{var}(X_2) + 2\text{cov}(X_1, X_2)
\]
Definition

Suppose $\mathbf{X} = (X_1, X_2, \ldots, X_N)$ is a vector of random variables. If $\mathbf{X}$ has pdf

$$f(\mathbf{x}) = (2\pi)^{-N/2} \det(\Sigma)^{-1/2} \exp \left( -\frac{1}{2}(\mathbf{x} - \mu)' \Sigma (\mathbf{x} - \mu) \right)$$

Then we will say $\mathbf{X}$ is a **Multivariate Normal Distribution**, $\mathbf{X} \sim \text{Multivariate Normal}(\mu, \Sigma)$

- Regularly used for likelihood, Bayesian, and other parametric inferences
Properties of the Multivariate Normal Distribution

Suppose $\mathbf{X} = (X_1, X_2, \ldots, X_N)$

$\mathbf{X} \sim \text{Multivariate Normal}(\mu, \Sigma)$

\[
\begin{align*}
E[\mathbf{X}] &= \mu \\
\text{cov}(\mathbf{X}) &= \Sigma
\end{align*}
\]
Properties of the Multivariate Normal Distribution

Suppose \( \mathbf{X} = (X_1, X_2, \ldots, X_N) \)
\( X \sim \text{Multivariate Normal}(\mu, \Sigma) \)

\[ E[\mathbf{X}] = \mu \]
\[ \text{cov}(\mathbf{X}) = \Sigma \]

So that,

\[
\Sigma = \begin{pmatrix}
\text{var}(X_1) & \text{cov}(X_1, X_2) & \cdots & \text{cov}(X_1, X_N) \\
\text{cov}(X_2, X_1) & \text{var}(X_2) & \cdots & \text{cov}(X_2, X_N) \\
\vdots & \vdots & \ddots & \vdots \\
\text{cov}(X_N, X_1) & \text{cov}(X_N, X_2) & \cdots & \text{var}(X_N)
\end{pmatrix}
\]
Multivariate Normal Distribution

Consider the (bivariate) special case where $\mu = (0, 0)$ and

$$\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
Multivariate Normal Distribution

Consider the (bivariate) special case where $\mu = (0, 0)$ and

$$
\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
$$

Then

$$f(x_1, x_2) = (2\pi)^{-2/2} 1^{-1/2} \exp \left( -\frac{1}{2} \left( (x - 0)' \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (x - 0) \right) \right) \Rightarrow \text{product of univariate standard normally distributed random variables}$$
Multivariate Normal Distribution

Consider the (bivariate) special case where $\mu = (0, 0)$ and

$$\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Then

$$f(x_1, x_2) = (2\pi)^{-2/2}1^{-1/2} \exp \left( -\frac{1}{2} \left( (x - 0)' \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (x - 0) \right) \right)$$

$$= \frac{1}{2\pi} \exp \left( -\frac{1}{2} (x_1^2 + x_2^2) \right) \Rightarrow \text{product of univariate standard normally distributed random variables}$$
Multivariate Normal Distribution

Consider the (bivariate) special case where $\boldsymbol{\mu} = (0, 0)$ and

$$
\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
$$

Then

$$
f(x_1, x_2) = (2\pi)^{-2/2} 1^{-1/2} \exp \left( -\frac{1}{2} \left( (x - 0)' \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (x - 0) \right) \right)
$$

$$
= \frac{1}{2\pi} \exp \left( -\frac{1}{2} (x_1^2 + x_2^2) \right)
$$

$$
= \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x_1^2}{2} \right) \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x_2^2}{2} \right)
$$

\(\Rightarrow\) product of univariate standard normally distributed random variables.
Multivariate Normal Distribution

Consider the (bivariate) special case where $\mathbf{\mu} = (0, 0)$ and

$$
\Sigma = \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}
$$

Then

$$
f(x_1, x_2) = (2\pi)^{-2/2}1^{-1/2} \exp \left( -\frac{1}{2} \left( (x - 0)' \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (x - 0) \right) \right)
$$

$$
= \frac{1}{2\pi} \exp \left( -\frac{1}{2} x_1^2 + x_2^2 \right)
$$

$$
= \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x_1^2}{2} \right) \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x_2^2}{2} \right)
$$

$\sim$ product of univariate standard normally distributed random variables
Standard Multivariate Normal

Definition
Suppose $Z = (Z_1, Z_2, \ldots, Z_N)$ is

$$Z \sim \text{Multivariate Normal}(0, I_N).$$

Then we will call $Z$ the standard multivariate normal.
Independence and Multivariate Normal

Proposition

Suppose $X$ and $Y$ are independent. Then

$$cov(X, Y) = 0$$
Proof.
Suppose $X$ and $Y$ are independent.
Proof.
Suppose $X$ and $Y$ are independent.

\[
\text{cov}(X, Y) = E[XY] - E[X]E[Y]
\]
Proof.
Suppose $X$ and $Y$ are independent.

$$\text{cov}(X, Y) = E[XY] - E[X]E[Y]$$

Calculating $E[XY]$
Proof.
Suppose $X$ and $Y$ are independent.

$$\text{cov}(X, Y) = E[XY] - E[X]E[Y]$$

Calculating $E[XY]$:

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) \, dx \, dy$$
Proof.
Suppose $X$ and $Y$ are independent.

$$\text{cov}(X, Y) = E[XY] - E[X]E[Y]$$

Calculating $E[XY]$

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) \, dx \, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x y f_X(x) f_Y(y) \, dx \, dy$$
Proof.

Suppose $X$ and $Y$ are independent.

$$\text{cov}(X, Y) = E[XY] - E[X]E[Y]$$

Calculating $E[XY]$

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)dx\,dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf_X(x)f_Y(y)dx\,dy$$

$$= \int_{-\infty}^{\infty} xf_X(x)dx \int_{-\infty}^{\infty} yf_Y(y)dy$$

Then $\text{cov}(X, Y) = 0$.

More generally if $X$ and $Y$ are independent, $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$ for functions $g: \mathbb{R} \to \mathbb{R}$ and $h: \mathbb{R} \to \mathbb{R}$.
Proof.
Suppose $X$ and $Y$ are independent.

\[
\text{cov}(X, Y) = E[XY] - E[X]E[Y]
\]

Calculating $E[XY]$

\[
E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)dxdy
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf_X(x)f_Y(y)dxdy
\]

\[
= \int_{-\infty}^{\infty} xf_X(x)dx \int_{-\infty}^{\infty} yf_Y(y)dy
\]

\[
= E[X]E[Y]
\]
Proof.
Suppose $X$ and $Y$ are independent.

\[
\text{cov}(X, Y) = E[XY] - E[X]E[Y]
\]

Calculating $E[XY]$

\[
E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) \, dx \, dy
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y) \, dx \, dy
\]

\[
= \int_{\infty}^{\infty} xf_X(x) \, dx \int_{-\infty}^{\infty} yf_Y(y) \, dy
\]

Then $\text{cov}(X, Y) = 0$. 
Proof.
Suppose $X$ and $Y$ are independent.

$$\text{cov}(X, Y) = E[XY] - E[X]E[Y]$$

Calculating $E[XY]$

$$E[XY] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) \, dx \, dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf_X(x)f_Y(y) \, dx \, dy$$

$$= \int_{-\infty}^{\infty} xf_X(x) \, dx \int_{-\infty}^{\infty} yf_Y(y) \, dy$$

$$= E[X]E[Y]$$

Then $\text{cov}(X, Y) = 0$.

- More generally if $X$ and $Y$ are independent,
  $$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$
  for functions $g : \mathbb{R} \rightarrow \mathbb{R}$ and $h : \mathbb{R} \rightarrow \mathbb{R}$. 
Zero covariance does not generally imply Independent

Suppose $X \in \{-1, 1\}$ with $P(X = 1) = P(X = -1) = 1/2$. Suppose $Y \in \{-1, 0, 1\}$ with $Y = 0$ if $X = -1$ and $P(Y = 1) = P(Y = -1)$ if $X = 1$.

$$E[XY] = \sum_{i \in \{-1,1\}} \sum_{j \in \{-1,0,1\}} ijP(X = i, Y = j)$$

$$= -1 \times 0 \times P(X = -1, Y = 0) + 1 \times 1 \times P(X = 1, Y = 1)$$

$$-1 \times 1 \times P(X = 1, Y = -1)$$

$$= 0 + P(X = 1, Y = 1) - P(X = 1, Y = -1)$$

$$= 0.25 - 0.25 = 0$$

$E[X] = 0$

$E[Y] = 0$
Proposition

Suppose \( \mathbf{X} \sim \text{Multivariate Normal}(\mu, \Sigma) \). where \( \mathbf{X} = (X_1, X_2, \ldots, X_N) \).

If \( \text{cov}(X_i, X_j) = 0 \), then \( X_i \) and \( X_j \) are independent.
Iterated Expectations

Proposition

Suppose $X$ and $Y$ are random variables. Then

$$E[X] = E[E[X|Y]]$$

- Inner Expectation is $E[X|Y] = \int_{-\infty}^{\infty} xf_{X|Y}(x|y) dx$.
- Outer expectation is over $y$.  

Justin Grimmer (Stanford University)
Iterated Expectations

Proof.

\[
\mathbb{E}[\mathbb{E}[X|Y]] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y}(x|y) f_Y(y) \, dx \, dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X|Y}(x|y) f_Y(y) \, dy \, dx = \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x,y) \, dy \, dx = \int_{-\infty}^{\infty} x f_X(x) \, dx = \mathbb{E}[X].
\]
Iterated Expectations

Proof.

\[
E[E[X|Y]] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_{X|Y}(x|y)f_Y(y)\,dx\,dy
\]
Iterated Expectations

Proof.

\[ E[E[X|Y]] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y}(x|y) f_Y(y) \, dx \, dy \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y}(x|y) f_Y(y) \, dy \, dx \]

\[ = \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_X(x, y) \, dy \, dx \]

\[ = E[X] \]
Iterated Expectations

Proof.

\[ E[E[X|Y]] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_{X|Y}(x|y)f_Y(y) \, dx \, dy \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_{X|Y}(x|y)f_Y(y) \, dy \, dx \]

\[ = \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x, y) \, dy \, dx \]
Iterated Expectations

Proof.

\[
E[E[X|Y]] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_{X|Y}(x|y)f_Y(y) \, dx \, dy
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_{X|Y}(x|y)f_Y(y) \, dy \, dx
\]

\[
= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x, y) \, dy \, dx
\]

\[
= \int_{-\infty}^{\infty} xf_X(x) \, dx
\]
Iterated Expectations

Proof.

\[
E[E[X|Y]] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_X|Y(x|y)f_Y(y) \, dx \, dy
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_X|Y(x|y)f_Y(y) \, dy \, dx
\]

\[
= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x, y) \, dy \, dx
\]

\[
= \int_{-\infty}^{\infty} xf_X(x) \, dx
\]

\[= E[X] \]
Iterated Expectations

Definition

Suppose $Y$ is a continuous random variable with $Y \in [0, 1]$ and pdf of $Y$ given by

$$f(y) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} y^{\alpha_1-1}(1-y)^{\alpha_2-1}$$

Then we will say $Y$ is a Beta distribution with parameters $\alpha_1$ and $\alpha_2$.

Equivalently,

$$Y \sim \text{Beta}(\alpha_1, \alpha_2)$$

- Beta is a distribution on proportions
- Beta is a special case of the Dirichlet distribution
- $E[Y] = \frac{\alpha_1}{\alpha_1 + \alpha_2}$
Iterated Expectations

Suppose

\[ \pi \sim \text{Beta}(\alpha_1, \alpha_2) \]
\[ Y|\pi, n \sim \text{Binomial}(n, \pi) \]

What is \( E[Y] \)?
Iterated Expectations

Suppose

\[ \pi \sim \text{Beta}(\alpha_1, \alpha_2) \]
\[ Y|\pi, n \sim \text{Binomial}(n, \pi) \]

What is \( E[Y] \)?

\[ E[Y] = E[E[Y|\pi]] \]
Iterated Expectations

Suppose

\[ \pi \sim \text{Beta}(\alpha_1, \alpha_2) \]
\[ Y | \pi, n \sim \text{Binomial}(n, \pi) \]

What is \( E[Y] \)?

\[
E[Y] = E[E[Y|\pi]] \\
= \int_{-\infty}^{\infty} \sum_{j=0}^{N} \binom{N}{j} j p(j|\pi) f(\pi) d\pi
\]
Iterated Expectations

Suppose

\[ \pi \sim \text{Beta}(\alpha_1, \alpha_2) \]

\[ Y|\pi, n \sim \text{Binomial}(n, \pi) \]

What is \( E[Y] \)?

\[
E[Y] = E[E[Y|\pi]]
\]

\[
= \int_{-\infty}^{\infty} \sum_{j=0}^{N} \binom{N}{j} j p(j|\pi) f(\pi) d\pi
\]

\[
= \int_{-\infty}^{\infty} N \pi f(\pi) d\pi
\]
Iterated Expectations

Suppose

\[ \pi \sim \text{Beta}(\alpha_1, \alpha_2) \]

\[ Y | \pi, n \sim \text{Binomial}(n, \pi) \]

What is \( E[Y] \)?

\[
E[Y] = E[E[Y|\pi]] \\
= \int_{-\infty}^{\infty} \sum_{j=0}^{N} \binom{N}{j} j p(j|\pi) f(\pi) d\pi \\
= \int_{-\infty}^{\infty} N \pi f(\pi) d\pi \\
= N \frac{\alpha_1}{\alpha_1 + \alpha_2}
\]
Proposition

Suppose $X$ is a random variable and $Y = g(X)$, where $g : \mathbb{R} \to \mathbb{R}$ that is a monotonic function.
Define $g^{-1} : \mathbb{R} \to \mathbb{R}$ such that $g^{-1}(g(X)) = X$ and is differentiable. Then,

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{\partial g^{-1}(y)}{\partial y} \right| \quad \text{if } y = g(x) \text{ for some } x$$

$$= 0 \quad \text{otherwise}$$
Change of Coordinates

Proof.

Suppose \( g(\cdot) \) is monotonically increasing (WLOG)
Change of Coordinates

Proof.
Suppose $g(\cdot)$ is monotonically increasing (WLOG)

$$F_Y(y) = P(Y \leq y)$$
Change of Coordinates

Proof.

Suppose $g(\cdot)$ is monotonically increasing (WLOG)

\[
F_Y(y) = P(Y \leq y) \\
= P(g(X) \leq y)
\]

Now differentiating to get the pdf

\[
\frac{\partial F_Y(y)}{\partial y} = \frac{\partial F_X(g^{-1}(y))}{\partial y} = f_X(g^{-1}(y)) \frac{\partial g^{-1}(Y)}{\partial y}
\]

Then this is a pdf because

\[
\frac{\partial g^{-1}(Y)}{\partial y} > 0.
\]
Change of Coordinates

Proof.

Suppose $g(\cdot)$ is monotonically increasing (WLOG)

\[
F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y))
\]

Now differentiating to get the pdf

\[
\frac{\partial F_Y(y)}{\partial y} = \frac{\partial F_X(g^{-1}(y))}{\partial y} = f_X(g^{-1}(y)) \frac{\partial g^{-1}(Y)}{\partial y}
\]

Then this is a pdf because $\frac{\partial g^{-1}(Y)}{\partial y} > 0$. 
Change of Coordinates

Proof.

Suppose $g(\cdot)$ is monotonically increasing (WLOG)

\[
F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))
\]
Change of Coordinates

Proof.
Suppose \( g(\cdot) \) is monotonically increasing (WLOG)

\[
F_Y(y) = P(Y \leq y) \\
= P(g(X) \leq y) \\
= P(X \leq g^{-1}(y)) \\
= F_X(g^{-1}(y))
\]

Now differentiating to get the pdf
Change of Coordinates

Proof.

Suppose $g(\cdot)$ is monotonically increasing (WLOG)

\[ F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y)) \]

Now differentiating to get the pdf

\[ \frac{\partial F_Y(y)}{\partial y} = \frac{\partial F_X(g^{-1}(y))}{\partial y} = f_X(g^{-1}(y)) \frac{\partial g^{-1}(Y)}{\partial y} \]
Change of Coordinates

Proof.

Suppose \( g(\cdot) \) is monotonically increasing (WLOG)

\[
F_Y(y) = P(Y \leq y) \\
= P(g(X) \leq y) \\
= P(X \leq g^{-1}(y)) \\
= F_X(g^{-1}(y))
\]

Now differentiating to get the pdf

\[
\frac{\partial F_Y(y)}{\partial y} = \frac{\partial F_X(g^{-1}(y))}{\partial y} \\
= f_X(g^{-1}(y)) \frac{\partial g^{-1}(Y)}{\partial y}
\]

Then this is a pdf because \( \frac{\partial g^{-1}(Y)}{\partial y} > 0. \)
Change of Coordinates

Suppose $X$ is a random variable with pdf $f_X(x)$. Suppose $Y = X^n$. Find $f_Y(y)$.
Suppose $X$ is a random variable with pdf $f_X(x)$. Suppose $Y = X^n$. Find $f_Y(y)$. Then $g^{-1}(x) = x^{1/n}$.
Suppose $X$ is a random variable with pdf $f_X(x)$. Suppose $Y = X^n$. Find $f_Y(y)$.
Then $g^{-1}(x) = x^{1/n}$.

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{\partial g^{-1}(Y)}{\partial y} \right|$$
Suppose $X$ is a random variable with pdf $f_X(x)$. Suppose $Y = X^n$. Find $f_Y(y)$.
Then $g^{-1}(x) = x^{1/n}$.

\[
f_Y(y) = f_X(g^{-1}(y)) \left| \frac{\partial g^{-1}(Y)}{\partial y} \right|
\]

\[
= f_X(y^{1/n}) \frac{y^{1/n-1}}{n}
\]
Change of Coordinates

Suppose \( X \) is a random variable with pdf \( f_X(x) \). Suppose \( Y = X^n \). Find \( f_Y(y) \).

Then \( g^{-1}(x) = x^{1/n} \).

\[
f_Y(y) = f_X(g^{-1}(y)) \left| \frac{\partial g^{-1}(Y)}{\partial y} \right|
\]

\[
= f_X(y^{1/n}) \frac{y^{1/n - 1}}{n}
\]

We’ve used this to derive many of the pdfs.
Moment Generating Functions

Definition

Suppose $X$ is a random variable with pdf $f$. Define,

$$E[X^n] = \int_{-\infty}^{\infty} x^n f(x) \, dx$$

We will call $X^n$ the $n^{th}$ moment of $X$

- By this definition $\text{var}(X) = \text{Second Moment} - \text{First Moment}^2$
- We are assuming that the integral converges
Proposition

Suppose $X$ is a random variable with pdf $f(x)$. Call $M(t) = E[e^{tX}]$,

$$M(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x) \, dx$$

We will call $M(t)$ the moment generating function, because:

$$\frac{\partial^n M(t)}{\partial t^n} \bigg|_0 = E[X^n]$$

(Assuming that we can interchange derivative and integral)
Moment Generating Functions

Proof.
Recall the Taylor Expansion of $e^{tX}$ at 0,
Moment Generating Functions

Proof.
Recall the Taylor Expansion of $e^{tx}$ at 0,

$$e^{tx} = 1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \ldots$$
Moment Generating Functions

Proof.
Recall the Taylor Expansion of $e^{tX}$ at 0,

$$e^{tX} = 1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \ldots$$

Then,
Moment Generating Functions

Proof.

Recall the Taylor Expansion of $e^{tx}$ at 0,

$$e^{tx} = 1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \ldots$$

Then,

$$E[e^{tx}] = 1 + tE[X] + \frac{t^2}{2!}E[X^2] + \frac{t^3}{3!}E[X^3] + \ldots$$
Moment Generating Functions

Proof.

Recall the Taylor Expansion of $e^{tX}$ at 0,

$$e^{tX} = 1 + tx + rac{t^2x^2}{2!} + rac{t^3x^3}{3!} + \ldots$$

Then,

$$E[e^{tX}] = 1 + tE[X] + \frac{t^2}{2!}E[X^2] + \frac{t^3}{3!}E[X^3] + \ldots$$

Differentiate once:
Moment Generating Functions

Proof.
Recall the Taylor Expansion of $e^{tX}$ at 0,

$$e^{tX} = 1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \ldots$$

Then,

$$E[e^{tX}] = 1 + tE[X] + \frac{t^2E[X^2]}{2!} + \frac{t^3E[X^3]}{3!} + \ldots$$

Differentiate once:

$$\frac{\partial M(t)}{\partial t} = 0 + E[X] + \frac{2t}{2!}E[X^2] + \ldots$$

$$M'(0) = 0 + E[X] + 0 + 0 + \ldots$$
Proof.
Differentiate $n$ times

Evaluated at 0, yields $M_n(0) = \mathbb{E}[X_n] - \frac{n!}{n!} \mathbb{E}[X_{n+1}] (n+1)! + \ldots$

If two random variables, $X$ and $Y$, have the same moment generating functions, then $F_X(x) = F_Y(y)$ for almost all $x$. 

Justin Grimmer (Stanford University)
Proof.

Differentiate \( n \) times

\[
\frac{\partial^n M(t)}{\partial^n t} = 0 + 0 + 0 + \ldots + \frac{n \times n - 1 \times \ldots \times 2 \times t^0 E[X^n]}{n!} + \frac{n! t E[X^{n+1}]}{(n+1)!} + \ldots
\]
Proof.

Differentiate $n$ times

\[
\frac{\partial^n M(t)}{\partial^n t} = 0 + 0 + 0 + \ldots + \frac{n \times n - 1 \times \ldots 2 \times t^0 E[X^n]}{n!} + \frac{n!tE[X^{n+1}]}{(n+1)!} + \ldots
\]

\[
= \frac{n!E[X^n]}{n!} + \frac{n!tE[X^{n+1}]}{(n+1)!} + \ldots
\]
Proof.

Differentiate \( n \) times

\[
\frac{\partial^n M(t)}{\partial^n t} = 0 + 0 + 0 + \ldots + \frac{n \times n - 1 \times \ldots 2 \times t^0 E[X^n]}{n!} + \frac{n!tE[X^{n+1}]}{(n+1)!} + \ldots
\]

\[
= \frac{n!E[X^n]}{n!} + \frac{n!tE[X^{n+1}]}{(n+1)!} + \ldots
\]

Evaluated at 0, yields \( M^n(0) = E[X^n] \)
Proof.
Differentiate $n$ times

$$\frac{\partial^n M(t)}{\partial^n t} = 0 + 0 + 0 + \ldots + \frac{n \times n - 1 \times \ldots 2 \times t^0 E[X^n]}{n!} + \frac{n! t E[X^{n+1}]}{(n + 1)!} + \ldots$$

$$= \frac{n! E[X^n]}{n!} + \frac{n! t E[X^{n+1}]}{(n + 1)!} + \ldots$$

Evaluated at 0, yields $M^n(0) = E[X^n]$
The Moments of the Normal Distribution

Suppose $Z \sim N(0, 1)$. 
The Moments of the Normal Distribution

Suppose $Z \sim N(0, 1)$. 

$$E[e^{tx}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2/2} dx$$
Suppose $Z \sim N(0, 1)$. 

\[
E[e^{tx}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2/2} \, dx
\]

\[
tx - \frac{1}{2}x^2 = -\frac{1}{2} ((x - t)^2 - t^2)
\]
The Moments of the Normal Distribution

Suppose $Z \sim N(0, 1)$.

\[ E[e^{tx}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx}e^{-x^2/2}dx \]

\[ tx - \frac{1}{2}x^2 = -\frac{1}{2}((x - t)^2 - t^2) \]

\[ E[e^{tx}] = \frac{1}{\sqrt{2\pi}} e^{t^2/2} \int_{-\infty}^{\infty} e^{-(x-t)^2/2}dx \]
The Moments of the Normal Distribution

Suppose $Z \sim N(0, 1)$.

$$E[e^{tX}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2/2} dx$$

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

$$E[e^{tX}] = \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} e^{-(x-t)^2/2} dx$$

$$= e^{\frac{t^2}{2}}$$
Extracting Moments of the Normal Distribution
Extracting Moments of the Normal Distribution

\[ M'(0) = E[X] = e^{t^2/2} t \big|_0 = 0 \]
Extracting Moments of the Normal Distribution

\[ M'(0) = E[X] = e^{t^2/2}t|_0 = 0 \]
\[ M''(0) = E[X^2] = e^{t^2/2}(t^2 + 1)|_0 = 1 \]
Extracting Moments of the Normal Distribution

\[
M'(0) = E[X] = e^{t^2/2} t |_{0} = 0
\]
\[
M''(0) = E[X^2] = e^{t^2/2} (t^2 + 1) |_{0} = 1
\]
\[
M'''(0) = E[X^3] = e^{t^2/2} t (t^2 + 3) |_{0} = 0
\]
Extracting Moments of the Normal Distribution

\[
\begin{align*}
M'(0) &= E[X] = e^{t^2/2}t|_0 = 0 \\
M''(0) &= E[X^2] = e^{t^2/2}(t^2 + 1)|_0 = 1 \\
M'''(0) &= E[X^3] = e^{t^2/2}t(t^2 + 3)|_0 = 0 \\
M''''(0) &= E[X^4] = e^{t^2/2}(t^4 + 6t^2 + 3)|_0 = 3
\end{align*}
\]
Extracting Moments of the Normal Distribution

\[ M'(0) = E[X] = e^{t^2/2} t |_0 = 0 \]
\[ M''(0) = E[X^2] = e^{t^2/2} (t^2 + 1) |_0 = 1 \]
\[ M'''(0) = E[X^3] = e^{t^2/2} t(t^2 + 3) |_0 = 0 \]
\[ M''''(0) = E[X^4] = e^{t^2/2} (t^4 + 6t^2 + 3) |_0 = 3 \]
\[ M^5(0) = E[X^5] = e^{t^2/2} t(t^4 + 10t^2 + 15) |_0 = 0 \]
Extracting Moments of the Normal Distribution

\[ M'(0) = E[X] = e^{t^2/2} t |_{0} = 0 \]
\[ M''(0) = E[X^2] = e^{t^2/2} (t^2 + 1) |_{0} = 1 \]
\[ M'''(0) = E[X^3] = e^{t^2/2} t(t^2 + 3) |_{0} = 0 \]
\[ M''''(0) = E[X^4] = e^{t^2/2} (t^4 + 6t^2 + 3) |_{0} = 3 \]
\[ M^5(0) = E[X^5] = e^{t^2/2} t(t^4 + 10t^2 + 15) |_{0} = 0 \]
\[ M^6(0) = E[X^6] = e^{t^2/2} (t^6 + 15t^4 + 45t^2 + 15) |_{0} = 15 \]
Extracting Moments of the Normal Distribution

\[
\begin{align*}
M'(0) &= E[X] = e^{t^2/2}t\big|_0 = 0 \\
M''(0) &= E[X^2] = e^{t^2/2}(t^2 + 1)\big|_0 = 1 \\
M'''(0) &= E[X^3] = e^{t^2/2}t(t^2 + 3)\big|_0 = 0 \\
M''''(0) &= E[X^4] = e^{t^2/2}(t^4 + 6t^2 + 3)\big|_0 = 3 \\
M^5(0) &= E[X^5] = e^{t^2/2}t(t^4 + 10t^2 + 15)\big|_0 = 0 \\
M^6(0) &= E[X^6] = e^{t^2/2}(t^6 + 15t^4 + 45t^2 + 15)\big|_0 = 15
\end{align*}
\]
Proposition

Suppose $X_i$ are a sequence of independent random variables. Define

$$Y = \sum_{i=1}^{N} X_i$$

Then

$$M_Y(t) = \prod_{i=1}^{N} M_{X_i}(t)$$
Proof.

\[ M_Y(t) = E[e^{tY}] \]
Proof.

\[ M_Y(t) = E[e^{tY}] = E[e^{t \sum_{i=1}^{N} X_i}] \]
Proof.

\[ M_Y(t) = E[e^{tY}] = E[e^{t \sum_{i=1}^{N} X_i}] = E[e^{tX_1 + tX_2 + \ldots + tX_N}] \]
Proof.

\[ M_Y(t) = E[e^{tY}] = E[e^{t \sum_{i=1}^{N} X_i}] = E[e^{tX_1 + tX_2 + \ldots + tX_N}] = E[e^{tX_1}]E[e^{tX_2}]\ldots E[e^{tX_N}] \text{ (by independence)} \]
Proof.

\[ M_Y(t) = E[e^{tY}] = E[e^{t \sum_{i=1}^{N} X_i}] = E[e^{tX_1 + tX_2 + \ldots + tX_N}] = E[e^{tX_1}]E[e^{tX_2}] \ldots E[e^{tX_N}] \text{ (by independence)} = \prod_{i=1}^{N} E[e^{tX_i}] \]
Tomorrow:
Sequences of Random Variables