Recap: Functions of multiple rv

- Given $X_1, \cdots, X_n$, random variables on the same probability space, $Z = g(X_1, \cdots, X_n)$ is a rv

Recap: Functions of multiple rv

- Given $X_1, \cdots, X_n$, random variables on the same probability space, $Z = g(X_1, \cdots, X_n)$ is a rv (if $g : \Re^n \to \Re$ is borel measurable).

Recap: Functions of multiple rv

- Given $X_1, \cdots, X_n$, random variables on the same probability space, $Z = g(X_1, \cdots, X_n)$ is a rv (if $g : \Re^n \to \Re$ is borel measurable).
- We can determine distribution of Z from the joint distribution of all X_i

Recap: Functions of multiple rv

- Given $X_1, \cdots, X_n$, random variables on the same probability space, $Z = g(X_1, \cdots, X_n)$ is a rv (if $g : \Re^n \to \Re$ is borel measurable).
- We can determine distribution of Z from the joint distribution of all X_i

$$F_Z(z) = P[Z \leq z] = P[g(X_1, \cdots, X_n) \leq z]$$

Recap: iid random variables

- $X_1, \cdots, X_n$ are said to be independent if events $[X_1 \in B_1], \cdots, [X_n \in B_n]$ are independent.

Recap: iid random variables

- $X_1, \cdots, X_n$ are said to be independent if events $[X_1 \in B_1], \cdots, [X_n \in B_n]$ are independent.
- If $X_1, \cdots, X_n$ are indepedent and all of them have the same distribution function then they are said to be iid – independent and identically distributed

Recap: Independence of functions of random variables

Recap: Independence of functions of random variables

- If $\mathbf{X}, \mathbf{Y}$ are vector random variables (or random vectors), independence implies $[\mathbf{X} \in B_1]$ is independent of $[\mathbf{Y} \in B_2]$ for all borel sets B_1, B_2 (in appropriate spaces).

Recap: Independence of functions of random variables

- If $\mathbf{X}, \mathbf{Y}$ are vector random variables (or random vectors), independence implies $[\mathbf{X} \in B_1]$ is independent of $[\mathbf{Y} \in B_2]$ for all borel sets B_1, B_2 (in appropriate spaces).
- Then $g(\mathbf{X})$ would be independent of $h(\mathbf{Y})$.

Recap: Independence of functions of random variables

- If $\mathbf{X}, \mathbf{Y}$ are vector random variables (or random vectors), independence implies $[\mathbf{X} \in B_1]$ is independent of $[\mathbf{Y} \in B_2]$ for all borel sets B_1, B_2 (in appropriate spaces).
- Then $g(\mathbf{X})$ would be independent of $h(\mathbf{Y})$.
- That is, suppose $X_1, \cdots, X_m, X_{m+1}, \cdots, X_{m+n}$ are independent.

Recap: Independence of functions of random variables

- If $\mathbf{X}, \mathbf{Y}$ are vector random variables (or random vectors), independence implies $[\mathbf{X} \in B_1]$ is independent of $[\mathbf{Y} \in B_2]$ for all borel sets B_1, B_2 (in appropriate spaces).
- Then $g(\mathbf{X})$ would be independent of $h(\mathbf{Y})$.
- That is, suppose $X_1, \cdots, X_m, X_{m+1}, \cdots, X_{m+n}$ are independent.
- Then, $g(X_1, \cdots, X_m)$ is independent of $h(X_{m+1}, \cdots, X_{m+n})$.

Recap: Max of a set of random variables

- Let $X_1, \cdots, X_n$ be independent and $Z = \max(X_1, \cdots, X_n)$

Recap: Max of a set of random variables

- Let $X_1, \cdots, X_n$ be independent and $Z = \max(X_1, \cdots, X_n)$

$$F_Z(z) \;=\; \prod_{i=1}^{n} F_{X_i}(z)$$

Recap: Max of a set of random variables

- Let $X_1, \cdots, X_n$ be independent and $Z = \max(X_1, \cdots, X_n)$

$$
\begin{aligned}
F_Z(z) &= \prod_{i=1}^{n} F_{X_i}(z) \\
&= (F(z))^n, \quad \text{if they are iid}
\end{aligned}
$$

Recap: Min of a set of random variables

- Let $X_1, \cdots, X_n$ be independent and $Z = \min(X_1, \cdots, X_n)$

Recap: Min of a set of random variables

- Let $X_1, \cdots, X_n$ be independent and $Z = \min(X_1, \cdots, X_n)$

$$F_Z(z) = 1 - \prod_{i=1}^{n} (1 - F_{X_i}(z))$$

Recap: Min of a set of random variables

- Let $X_1, \cdots, X_n$ be independent and $Z = \min(X_1, \cdots, X_n)$

$$
\begin{aligned}
F_Z(z) &= 1 - \prod_{i=1}^{n} (1 - F_{X_i}(z)) \\
&= 1 - (1 - F(z))^n, \quad \text{if they are iid}
\end{aligned}
$$

Recap: Order Statistics

- Let $X_1, \cdots, X_n$ be iid with density f.

Recap: Order Statistics

- Let $X_1, \cdots, X_n$ be iid with density f.
- Let $X_{(k)}$ denote the k^{th} smallest of these.

Recap: Order Statistics

- Let $X_1, \cdots, X_n$ be iid with density f.
- Let $X_{(k)}$ denote the k^{th} smallest of these.
- $X_{(1)} = \min(X_1, \cdots, X_n), \quad X_{(n)} = \max(X_1, \cdots, X_n)$

Recap: Order Statistics

- Let $X_1, \cdots, X_n$ be iid with density f.
- Let $X_{(k)}$ denote the k^{th} smallest of these.
- $X_{(1)} = \min(X_1, \cdots, X_n), \quad X_{(n)} = \max(X_1, \cdots, X_n)$
- We have:

$$F_{X_{(k)}}(y) = \sum_{j=k}^{n} {}^nC_j (F(y))^j (1 - F(y))^{n-j}$$

Recap: Order Statistics

- Let $X_1, \cdots, X_n$ be iid with density f.
- Let $X_{(k)}$ denote the k^{th} smallest of these.
- $X_{(1)} = \min(X_1, \cdots, X_n), \quad X_{(n)} = \max(X_1, \cdots, X_n)$
- We have:

$$F_{X_{(k)}}(y) = \sum_{j=k}^{n} {}^{n}C_j (F(y))^j (1 - F(y))^{n-j}$$

- Joint distribution of $X_{(1)}, \cdots X_{(n)}$ is called the order statistics.

$$f_{X_{(1)} \cdots X_{(n)}}(x_1, \cdots x_n) = n! \prod_{i=1}^{n} f(x_i), \;\; x_1 < x_2 < \cdots < x_n$$

Recap: Sum of two independent rv

- Let X, Y be random variables and $Z = X + Y$.

Recap: Sum of two independent rv

- Let X, Y be random variables and $Z = X + Y$.
- If X, Y are discrete

$$f_Z(z) \;=\; \sum_k f_{XY}(k, z-k)$$

Recap: Sum of two independent rv

- Let X, Y be random variables and $Z = X + Y$.
- If X, Y are discrete

$$\begin{aligned} f_Z(z) &= \sum_k f_{XY}(k, z-k) \\ &= \sum_k f_X(k) f_Y(z-k) \quad \text{when } X, Y \text{ are independent.} \end{aligned}$$

Recap: Sum of two independent rv

- Let X, Y be random variables and $Z = X + Y$.
- If X, Y are discrete

$$\begin{aligned} f_Z(z) &= \sum_k f_{XY}(k, z-k) \\ &= \sum_k f_X(k) f_Y(z-k) \ \text{ when } X, Y \text{ are independent.} \end{aligned}$$

- If X, Y have a joint density

$$f_Z(z) = \int_{-\infty}^{\infty} f_{XY}(t, z-t)\, dt$$

Recap: Sum of two independent rv

- Let X, Y be random variables and $Z = X + Y$.
- If X, Y are discrete

$$\begin{aligned} f_Z(z) &= \sum_k f_{XY}(k, z-k) \\ &= \sum_k f_X(k) f_Y(z-k) \;\; \text{when } X, Y \text{ are independent.} \end{aligned}$$

- If X, Y have a joint density

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\, dt \\ &= \int_{-\infty}^{\infty} f_X(t)\, f_Y(z-t)\, dt, \;\; \text{when } X, Y \text{ are independe} \end{aligned}$$

Recap: Sum of two independent rv

- Let X, Y be random variables and $Z = X + Y$.
- If X, Y are discrete

$$
\begin{aligned}
f_Z(z) &= \sum_k f_{XY}(k, z-k) \\
&= \sum_k f_X(k) f_Y(z-k) \quad \text{when } X, Y \text{ are independent.}
\end{aligned}
$$

- If X, Y have a joint density

$$
\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\, dt \\
&= \int_{-\infty}^{\infty} f_X(t)\, f_Y(z-t)\, dt, \quad \text{when } X, Y \text{ are independe}
\end{aligned}
$$

Density of sum of independent random variables is the convolution of their densities.

Recap: sum of two independent random variables

- Given independent random variables X, Y

Recap: sum of two independent random variables

- Given independent random variables X, Y
- If $X \sim Poisson(\lambda_1)$ and $Y \sim Poisson(\lambda_2)$
  $X + Y \sim Poisson(\lambda_1 + \lambda_2)$

Recap: sum of two independent random variables

- Given independent random variables X, Y
- If $X \sim Poisson(\lambda_1)$ and $Y \sim Poisson(\lambda_2)$
  $X + Y \sim Poisson(\lambda_1 + \lambda_2)$
- If $X \sim Gamma(\alpha_1, \lambda)$ and $Y \sim Gamma(\alpha_2, \lambda)$
  $X + Y \sim Gamma(\alpha_1 + \alpha_2, \lambda)$

Recap: sum of two independent random variables

- Given independent random variables X, Y
- If $X \sim Poisson(\lambda_1)$ and $Y \sim Poisson(\lambda_2)$
  $X + Y \sim Poisson(\lambda_1 + \lambda_2)$
- If $X \sim Gamma(\alpha_1, \lambda)$ and $Y \sim Gamma(\alpha_2, \lambda)$
  $X + Y \sim Gamma(\alpha_1 + \alpha_2, \lambda)$
- If $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$
  $X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Recap: Theorem for functions of cont rv

- $X_1, \cdots X_n$ are continuous rv with joint density

Recap: Theorem for functions of cont rv

- $X_1, \cdots X_n$ are continuous rv with joint density

$$Y_1 = g_1(X_1, \cdots, X_n) \quad \cdots \quad Y_n = g_n(X_1, \cdots, X_n)$$

Recap: Theorem for functions of cont rv

- $X_1, \cdots X_n$ are continuous rv with joint density

$$Y_1 = g_1(X_1, \cdots, X_n) \quad \cdots \quad Y_n = g_n(X_1, \cdots, X_n)$$

- The transformation is continuous with continuous first partials and is invertible and

$$X_1 = h_1(Y_1, \cdots, Y_n) \quad \cdots \quad X_n = h_n(Y_1, \cdots, Y_n)$$

Recap: Theorem for functions of cont rv

- $X_1, \cdots X_n$ are continuous rv with joint density

$$Y_1 = g_1(X_1, \cdots, X_n) \quad \cdots \quad Y_n = g_n(X_1, \cdots, X_n)$$

- The transformation is continuous with continuous first partials and is invertible and

$$X_1 = h_1(Y_1, \cdots, Y_n) \quad \cdots \quad X_n = h_n(Y_1, \cdots, Y_n)$$

- Jacobian of the inverse transform, J, is non-zero

Recap: Theorem for functions of cont rv

- $X_1, \cdots X_n$ are continuous rv with joint density

$$Y_1 = g_1(X_1, \cdots, X_n) \quad \cdots \quad Y_n = g_n(X_1, \cdots, X_n)$$

- The transformation is continuous with continuous first partials and is invertible and

$$X_1 = h_1(Y_1, \cdots, Y_n) \quad \cdots \quad X_n = h_n(Y_1, \cdots, Y_n)$$

- Jacobian of the inverse transform, J, is non-zero
- Then the density of $\mathbf{Y}$ is

$$f_{Y_1 \cdots Y_n}(y_1, \cdots, y_n) = |J| f_{X_1 \cdots X_n}\left(h_1(y_1, \cdots, y_n), \cdots, h_n(y_1, \cdots, y_n)\right)$$

Recap: Theorem for functions of cont rv

- $X_1, \cdots X_n$ are continuous rv with joint density

$$Y_1 = g_1(X_1, \cdots, X_n) \quad \cdots \quad Y_n = g_n(X_1, \cdots, X_n)$$

- The transformation is continuous with continuous first partials and is invertible and

$$X_1 = h_1(Y_1, \cdots, Y_n) \quad \cdots \quad X_n = h_n(Y_1, \cdots, Y_n)$$

- Jacobian of the inverse transform, J, is non-zero
- Then the density of $\mathbf{Y}$ is

$$f_{Y_1 \cdots Y_n}(y_1, \cdots, y_n) = |J| f_{X_1 \cdots X_n}\left(h_1(y_1, \cdots, y_n), \cdots, h_n(y_1, \cdots, y_n)\right)$$

- Called multidimensional change of variable formula

- Let X, Y have joint density f_{XY}. Let $Z = X + Y$.

- Let X, Y have joint density f_{XY}. Let $Z = X + Y$.
- We want to find f_Z using the theorem.

- Let X, Y have joint density f_{XY}. Let $Z = X + Y$.
- We want to find f_Z using the theorem.
- To use the theorem, we need an invertible transformation of $\Re^2$ onto $\Re^2$ of which one component is $x + y$.

- Let X, Y have joint density f_{XY}. Let $Z = X + Y$.
- We want to find f_Z using the theorem.
- To use the theorem, we need an invertible transformation of $\Re^2$ onto $\Re^2$ of which one component is $x + y$.
- Take $Z = X + Y$ and $W = X - Y$. This is invertible.

- Let X, Y have joint density f_{XY}. Let $Z = X + Y$.
- We want to find f_Z using the theorem.
- To use the theorem, we need an invertible transformation of $\Re^2$ onto $\Re^2$ of which one component is $x + y$.
- Take $Z = X + Y$ and $W = X - Y$. This is invertible.
- $X = (Z + W)/2$ and $Y = (Z - W)/2$.

- Let X, Y have joint density f_{XY}. Let $Z = X + Y$.
- We want to find f_Z using the theorem.
- To use the theorem, we need an invertible transformation of $\Re^2$ onto $\Re^2$ of which one component is $x + y$.
- Take $Z = X + Y$ and $W = X - Y$. This is invertible.
- $X = (Z + W)/2$ and $Y = (Z - W)/2$. The Jacobian is

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

- Let X, Y have joint density f_{XY}. Let $Z = X + Y$.
- We want to find f_Z using the theorem.
- To use the theorem, we need an invertible transformation of $\Re^2$ onto $\Re^2$ of which one component is $x + y$.
- Take $Z = X + Y$ and $W = X - Y$. This is invertible.
- $X = (Z + W)/2$ and $Y = (Z - W)/2$. The Jacobian is

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

- Hence we get

$$f_{ZW}(z, w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- Let X, Y have joint density f_{XY}. Let $Z = X + Y$.
- We want to find f_Z using the theorem.
- To use the theorem, we need an invertible transformation of $\Re^2$ onto $\Re^2$ of which one component is $x + y$.
- Take $Z = X + Y$ and $W = X - Y$. This is invertible.
- $X = (Z + W)/2$ and $Y = (Z - W)/2$. The Jacobian is

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

- Hence we get

$$f_{ZW}(z, w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- Now we get density of Z as

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right) \, dw$$

- let $Z = X + Y$ and $W = X - Y$. Then

- let $Z = X + Y$ and $W = X - Y$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)\, dw$$

- let $Z = X + Y$ and $W = X - Y$. Then

$$f_Z(z) \;=\; \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)\, dw$$

$$\text{change the variable: } \; t = \frac{z+w}{2}$$

- let $Z = X + Y$ and $W = X - Y$. Then

$$f_Z(z) \;=\; \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)\, dw$$

change the variable: $t = \dfrac{z+w}{2} \quad \Rightarrow dt = \dfrac{1}{2}\, dw$

- let $Z = X + Y$ and $W = X - Y$. Then

$$f_Z(z) \;=\; \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)\, dw$$

change the variable: $t = \dfrac{z+w}{2} \quad \Rightarrow dt = \dfrac{1}{2}\, dw$

$$\Rightarrow \quad w = 2t - z \quad \Rightarrow z - w = 2z - 2t$$

- let $Z = X + Y$ and $W = X - Y$. Then

$$
\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)\, dw \\
& \qquad \text{change the variable: } t = \frac{z+w}{2} \quad \Rightarrow dt = \frac{1}{2}\, dw \\
& \qquad\qquad \Rightarrow \quad w = 2t - z \quad \Rightarrow z - w = 2z - 2t \\
f_Z(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\, dt
\end{aligned}
$$

- let $Z = X + Y$ and $W = X - Y$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right) \, dw$$

change the variable: $t = \dfrac{z+w}{2} \quad \Rightarrow dt = \dfrac{1}{2}\, dw$

$$\Rightarrow \quad w = 2t - z \quad \Rightarrow z - w = 2z - 2t$$

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t) \, dt \\ &= \int_{-\infty}^{\infty} f_{XY}(z-s, s) \, ds, \end{aligned}$$

- let $Z = X + Y$ and $W = X - Y$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)\, dw$$

change the variable: $t = \dfrac{z+w}{2} \quad \Rightarrow dt = \dfrac{1}{2}\, dw$

$$\Rightarrow \quad w = 2t - z \quad \Rightarrow z - w = 2z - 2t$$

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\, dt \\ &= \int_{-\infty}^{\infty} f_{XY}(z-s, s)\, ds, \end{aligned}$$

- We get same result as earlier.

- let $Z = X + Y$ and $W = X - Y$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)\, dw$$

change the variable: $t = \dfrac{z+w}{2} \quad \Rightarrow dt = \dfrac{1}{2}\, dw$

$$\Rightarrow \quad w = 2t - z \quad \Rightarrow z - w = 2z - 2t$$

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\, dt \\ &= \int_{-\infty}^{\infty} f_{XY}(z-s, s)\, ds, \end{aligned}$$

- We get same result as earlier. If, X, Y are independent

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\, f_Y(z-t)\, dt$$

- let $Z = X + Y$ and $W = X - Y$. We got

- let $Z = X + Y$ and $W = X - Y$. We got

$$f_{ZW}(z,w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- let $Z = X + Y$ and $W = X - Y$. We got

$$f_{ZW}(z, w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- Now we can calculate f_W also.

- let $Z = X + Y$ and $W = X - Y$. We got

$$f_{ZW}(z,w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- Now we can calculate f_W also.

$$f_W(w) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)\, dz$$

- let $Z = X + Y$ and $W = X - Y$. We got

$$f_{ZW}(z,w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- Now we can calculate f_W also.

$$f_W(w) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right) \, dz$$

change the variable: $t = \dfrac{z+w}{2}$

- let $Z = X + Y$ and $W = X - Y$. We got

$$f_{ZW}(z, w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- Now we can calculate f_W also.

$$f_W(w) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right) dz$$

$$\text{change the variable: } t = \frac{z+w}{2} \quad \Rightarrow dt = \frac{1}{2}\, dz$$

- let $Z = X + Y$ and $W = X - Y$. We got

$$f_{ZW}(z, w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- Now we can calculate f_W also.

$$f_W(w) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right) \, dz$$

change the variable: $t = \dfrac{z+w}{2} \quad \Rightarrow dt = \dfrac{1}{2}\, dz$

$\Rightarrow \quad z = 2t - w \quad \Rightarrow z - w = 2t - 2w$

- let $Z = X + Y$ and $W = X - Y$. We got

$$f_{ZW}(z, w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- Now we can calculate f_W also.

$$f_W(w) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right) \, dz$$

change the variable: $t = \dfrac{z+w}{2} \quad \Rightarrow dt = \dfrac{1}{2}\, dz$

$\Rightarrow \quad z = 2t - w \quad \Rightarrow z - w = 2t - 2w$

$$f_W(w) = \int_{-\infty}^{\infty} f_{XY}(t, t-w) \, dt$$

- let $Z = X + Y$ and $W = X - Y$. We got

$$f_{ZW}(z, w) = \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right)$$

- Now we can calculate f_W also.

$$f_W(w) = \int_{-\infty}^{\infty} \frac{1}{2} f_{XY}\left(\frac{z+w}{2}, \frac{z-w}{2}\right) dz$$

$$\text{change the variable: } t = \frac{z+w}{2} \quad \Rightarrow dt = \frac{1}{2}\, dz$$

$$\Rightarrow \quad z = 2t - w \quad \Rightarrow z - w = 2t - 2w$$

$$f_W(w) = \int_{-\infty}^{\infty} f_{XY}(t, t-w)\, dt$$

$$= \int_{-\infty}^{\infty} f_{XY}(s+w, s) ds,$$

Example

- Let X, Y be iid $U[0,\ 1]$. Let $Z = X - Y$.

Example

- Let X, Y be iid $U[0,\ 1]$. Let $Z = X - Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\ f_Y(t-z)\ dt$$

Example

- Let X, Y be iid $U[0,\ 1]$. Let $Z = X - Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\ f_Y(t-z)\ dt$$

- For the integrand to be non-zero

Example

- Let X, Y be iid $U[0,\ 1]$. Let $Z = X - Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\ f_Y(t-z)\ dt$$

- For the integrand to be non-zero
  - $0 \leq t \leq 1 \ \Rightarrow \ t \geq 0, \ \ t \leq 1$

Example

- Let X, Y be iid $U[0,\ 1]$. Let $Z = X - Y$.
$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\ f_Y(t-z)\ dt$$
- For the integrand to be non-zero
  - $0 \leq t \leq 1 \ \Rightarrow \ t \geq 0, \ \ t \leq 1$
  - $0 \leq t - z \leq 1 \ \Rightarrow \ t \geq z, \ \ t \leq 1 + z$

Example

- Let X, Y be iid $U[0,\ 1]$. Let $Z = X - Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\ f_Y(t-z)\ dt$$

- For the integrand to be non-zero
  - $0 \le t \le 1 \ \Rightarrow \ t \ge 0,\ \ t \le 1$
  - $0 \le t - z \le 1 \ \Rightarrow \ t \ge z,\ \ t \le 1 + z$
  - $\Rightarrow \quad \max(0,\ z) \le t \le \min(1,\ 1+z)$

Example

- Let X, Y be iid $U[0,\ 1]$. Let $Z = X - Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\ f_Y(t - z)\ dt$$

- For the integrand to be non-zero
  - $0 \leq t \leq 1 \ \Rightarrow \ t \geq 0,\ \ t \leq 1$
  - $0 \leq t - z \leq 1 \ \Rightarrow \ t \geq z,\ \ t \leq 1 + z$
  - $\Rightarrow \quad \max(0,\ z) \leq t \leq \min(1,\ 1 + z)$
- Thus, we get density as (note $Z \in (-1, 1)$)

$$f_Z(z) = \begin{cases} \int_0^{1+z} 1\ dt = 1 + z, & \text{if } -1 \leq z \leq 0 \\ & \end{cases}$$

Example

- Let X, Y be iid $U[0,\ 1]$. Let $Z = X - Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\ f_Y(t-z)\ dt$$

- For the integrand to be non-zero
  - $0 \leq t \leq 1 \ \Rightarrow \ t \geq 0,\ \ t \leq 1$
  - $0 \leq t - z \leq 1 \ \Rightarrow \ t \geq z,\ \ t \leq 1 + z$
  - $\Rightarrow \ \max(0,\ z) \leq t \leq \min(1,\ 1+z)$
- Thus, we get density as (note $Z \in (-1, 1)$)

$$f_Z(z) = \begin{cases} \int_0^{1+z} 1\ dt = 1 + z, & \text{if } -1 \leq z \leq 0 \\ \int_z^1 1\ dt = 1 - z, & 0 \leq z \leq 1 \end{cases}$$

Example

- Let X, Y be iid $U[0,\ 1]$. Let $Z = X - Y$.

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\ f_Y(t-z)\ dt$$

- For the integrand to be non-zero
  - $0 \leq t \leq 1 \quad \Rightarrow \quad t \geq 0,\ \ t \leq 1$
  - $0 \leq t - z \leq 1 \quad \Rightarrow \quad t \geq z,\ \ t \leq 1 + z$
  - $\Rightarrow \quad \max(0,\ z) \leq t \leq \min(1,\ 1+z)$
- Thus, we get density as (note $Z \in (-1, 1)$)

$$f_Z(z) = \begin{cases} \int_0^{1+z} 1\ dt = 1 + z, & \text{if } -1 \leq z \leq 0 \\ \int_z^1 1\ dt = 1 - z, & 0 \leq z \leq 1 \end{cases}$$

- Thus, when $X, Y \sim U(0, 1)$ iid

$$f_{X-Y}(z) = 1 - |z|, \quad -1 < z < 1$$

- We showed that

- We showed that

$$f_{X+Y}(z) \;=\; \int_{-\infty}^{\infty} f_{XY}(t, z-t)\; dt = \int_{-\infty}^{\infty} f_{XY}(z-t, t)\; dt$$

- We showed that

$$
\begin{aligned}
f_{X+Y}(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\ dt = \int_{-\infty}^{\infty} f_{XY}(z-t, t)\ dt \\
f_{X-Y}(w) &= \int_{-\infty}^{\infty} f_{XY}(t, t-w)\ dt = \int_{-\infty}^{\infty} f_{XY}(t+w, t) dt
\end{aligned}
$$

- We showed that

$$\begin{aligned} f_{X+Y}(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\; dt = \int_{-\infty}^{\infty} f_{XY}(z-t, t)\; dt \\ f_{X-Y}(w) &= \int_{-\infty}^{\infty} f_{XY}(t, t-w)\; dt = \int_{-\infty}^{\infty} f_{XY}(t+w, t) dt \end{aligned}$$

- Suppose X, Y are discrete. Then we have

- We showed that

$$
\begin{aligned}
f_{X+Y}(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\, dt = \int_{-\infty}^{\infty} f_{XY}(z-t, t)\, dt \\
f_{X-Y}(w) &= \int_{-\infty}^{\infty} f_{XY}(t, t-w)\, dt = \int_{-\infty}^{\infty} f_{XY}(t+w, t) dt
\end{aligned}
$$

- Suppose X, Y are discrete. Then we have

$$
f_{X+Y}(z) = P[X+Y=z] = \sum_k P[X=k, Y=z-k]
$$

- We showed that

$$\begin{aligned} f_{X+Y}(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\ dt = \int_{-\infty}^{\infty} f_{XY}(z-t, t)\ dt \\ f_{X-Y}(w) &= \int_{-\infty}^{\infty} f_{XY}(t, t-w)\ dt = \int_{-\infty}^{\infty} f_{XY}(t+w, t) dt \end{aligned}$$

- Suppose X, Y are discrete. Then we have

$$\begin{aligned} f_{X+Y}(z) &= P[X+Y=z] = \sum_k P[X=k, Y=z-k] \\ &= \sum_k f_{XY}(k, z-k) \end{aligned}$$

- We showed that

$$
\begin{aligned}
f_{X+Y}(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\ dt = \int_{-\infty}^{\infty} f_{XY}(z-t, t)\ dt \\
f_{X-Y}(w) &= \int_{-\infty}^{\infty} f_{XY}(t, t-w)\ dt = \int_{-\infty}^{\infty} f_{XY}(t+w, t) dt
\end{aligned}
$$

- Suppose X, Y are discrete. Then we have

$$
\begin{aligned}
f_{X+Y}(z) &= P[X+Y=z] = \sum_k P[X=k, Y=z-k] \\
&= \sum_k f_{XY}(k, z-k) \\
f_{X-Y}(w) &= P[X-Y=w] = \sum_k P[X=k, Y=k-w]
\end{aligned}
$$

- We showed that

$$
\begin{aligned}
f_{X+Y}(z) &= \int_{-\infty}^{\infty} f_{XY}(t, z-t)\; dt = \int_{-\infty}^{\infty} f_{XY}(z-t, t)\; dt \\
f_{X-Y}(w) &= \int_{-\infty}^{\infty} f_{XY}(t, t-w)\; dt = \int_{-\infty}^{\infty} f_{XY}(t+w, t) dt
\end{aligned}
$$

- Suppose X, Y are discrete. Then we have

$$
\begin{aligned}
f_{X+Y}(z) &= P[X+Y=z] = \sum_k P[X=k, Y=z-k] \\
&= \sum_k f_{XY}(k, z-k) \\
f_{X-Y}(w) &= P[X-Y=w] = \sum_k P[X=k, Y=k-w] \\
&= \sum_k f_{XY}(k, k-w)
\end{aligned}
$$

Distribution of product of random variables

- We want density of $Z = XY$.

Distribution of product of random variables

- We want density of $Z = XY$.
- We need one more function to make an invertible transformation

Distribution of product of random variables

- We want density of $Z = XY$.
- We need one more function to make an invertible transformation
- A possible choice: $Z = XY \quad W = Y$

Distribution of product of random variables

- We want density of $Z = XY$.
- We need one more function to make an invertible transformation
- A possible choice: $Z = XY \quad W = Y$
- This is invertible: $X = Z/W \quad Y = W$

Distribution of product of random variables

- We want density of $Z = XY$.
- We need one more function to make an invertible transformation
- A possible choice: $Z = XY \quad W = Y$
- This is invertible: $X = Z/W \quad Y = W$

$$J = \begin{vmatrix} \frac{1}{w} & \frac{-z}{w^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{w}$$

Distribution of product of random variables

- We want density of $Z = XY$.
- We need one more function to make an invertible transformation
- A possible choice: $Z = XY \quad W = Y$
- This is invertible: $X = Z/W \quad Y = W$

$$J = \begin{vmatrix} \frac{1}{w} & \frac{-z}{w^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{w}$$

- Hence we get

$$f_{ZW}(z,w) = \left|\frac{1}{w}\right| \, f_{XY}\left(\frac{z}{w}, w\right)$$

Distribution of product of random variables

- We want density of $Z = XY$.
- We need one more function to make an invertible transformation
- A possible choice: $Z = XY \quad W = Y$
- This is invertible: $X = Z/W \quad Y = W$

$$J = \begin{vmatrix} \frac{1}{w} & \frac{-z}{w^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{w}$$

- Hence we get

$$f_{ZW}(z, w) = \left|\frac{1}{w}\right| \, f_{XY}\left(\frac{z}{w}, w\right)$$

- Thus we get the density of product as

$$f_Z(z) = \int_{-\infty}^{\infty} \left|\frac{1}{w}\right| \, f_{XY}\left(\frac{z}{w}, w\right) \, dw$$

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = XY$.

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = XY$.

$$f_Z(z) = \int_{-\infty}^{\infty} \left|\frac{1}{w}\right| \ f_X\left(\frac{z}{w}\right) \ f_Y(w) \ dw$$

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = XY$.

$$f_Z(z) = \int_{-\infty}^{\infty} \left|\frac{1}{w}\right| \ f_X\left(\frac{z}{w}\right) \ f_Y(w) \ dw$$

- We need: $0 < w < 1$ and $0 < \frac{z}{w} < 1$. Hence

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = XY$.

$$f_Z(z) = \int_{-\infty}^{\infty} \left|\frac{1}{w}\right| \ f_X\left(\frac{z}{w}\right) \ f_Y(w) \ dw$$

- We need: $0 < w < 1$ and $0 < \frac{z}{w} < 1$. Hence

$$f_Z(z) = \int_z^1 \left|\frac{1}{w}\right| \ dw$$

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = XY$.

$$f_Z(z) = \int_{-\infty}^{\infty} \left|\frac{1}{w}\right| \ f_X\left(\frac{z}{w}\right) \ f_Y(w) \ dw$$

- We need: $0 < w < 1$ and $0 < \frac{z}{w} < 1$. Hence

$$f_Z(z) = \int_z^1 \left|\frac{1}{w}\right| \ dw = \int_z^1 \frac{1}{w} \ dw = -\ln(z), \quad 0 < z < 1$$

- X, Y have joint density and $Z = XY$. Then

- X, Y have joint density and $Z = XY$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} \left|\frac{1}{w}\right| f_{XY}\left(\frac{z}{w}.w\right) dw$$

- X, Y have joint density and $Z = XY$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} \left| \frac{1}{w} \right| \; f_{XY}\left(\frac{z}{w}.w \right) \; dw$$

Suppose X, Y are discrete and $Z = XY$

- X, Y have joint density and $Z = XY$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} \left|\frac{1}{w}\right| \; f_{XY}\left(\frac{z}{w}.w\right) \; dw$$

Suppose X, Y are discrete and $Z = XY$

$$f_Z(0) \;\; = \;\; P[X = 0 \text{ or } Y = 0] = \sum_x f_{XY}(x, 0) + \sum_y f_{XY}(0, y)$$

- X, Y have joint density and $Z = XY$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} \left|\frac{1}{w}\right| \; f_{XY}\left(\frac{z}{w}.w\right) \; dw$$

Suppose X, Y are discrete and $Z = XY$

$$\begin{aligned} f_Z(0) &= P[X = 0 \text{ or } Y = 0] = \sum_x f_{XY}(x, 0) + \sum_y f_{XY}(0, y) \\ f_Z(k) &= \sum_{y \neq 0} P\left[X = \frac{k}{y}, Y = y\right] = \sum_{y \neq 0} f_{XY}\left(\frac{k}{y}, y\right), \; k \neq 0 \end{aligned}$$

- X, Y have joint density and $Z = XY$. Then

$$f_Z(z) = \int_{-\infty}^{\infty} \left| \frac{1}{w} \right| \, f_{XY}\left(\frac{z}{w}.w\right) \, dw$$

Suppose X, Y are discrete and $Z = XY$

$$\begin{aligned} f_Z(0) &= P[X = 0 \text{ or } Y = 0] = \sum_x f_{XY}(x, 0) + \sum_y f_{XY}(0, y) \\ f_Z(k) &= \sum_{y \neq 0} P\left[X = \frac{k}{y}, Y = y\right] = \sum_{y \neq 0} f_{XY}\left(\frac{k}{y}, y\right), \; k \neq 0 \end{aligned}$$

- We cannot always interchange density and mass functions!!

- We wanted density of $Z = XY$.

- We wanted density of $Z = XY$.
- We used: $Z = XY$ and $W = Y$.

- We wanted density of $Z = XY$.
- We used: $Z = XY$ and $W = Y$.
- We could have used: $Z = XY$ and $W = X$.

- We wanted density of $Z = XY$.
- We used: $Z = XY$ and $W = Y$.
- We could have used: $Z = XY$ and $W = X$.
- This is invertible: $X = W$ and $Y = Z/W$.

- We wanted density of $Z = XY$.
- We used: $Z = XY$ and $W = Y$.
- We could have used: $Z = XY$ and $W = X$.
- This is invertible: $X = W$ and $Y = Z/W$.

$$J = \begin{vmatrix} 0 & 1 \\ \frac{1}{w} & \frac{-z}{w^2} \end{vmatrix} = -\frac{1}{w}$$

- We wanted density of $Z = XY$.
- We used: $Z = XY$ and $W = Y$.
- We could have used: $Z = XY$ and $W = X$.
- This is invertible: $X = W$ and $Y = Z/W$.

$$J = \begin{vmatrix} 0 & 1 \\ \frac{1}{w} & \frac{-z}{w^2} \end{vmatrix} = -\frac{1}{w}$$

- This gives

- We wanted density of $Z = XY$.
- We used: $Z = XY$ and $W = Y$.
- We could have used: $Z = XY$ and $W = X$.
- This is invertible: $X = W$ and $Y = Z/W$.

$$J = \begin{vmatrix} 0 & 1 \\ \frac{1}{w} & \frac{-z}{w^2} \end{vmatrix} = -\frac{1}{w}$$

- This gives

$$f_{ZW}(z, w) = \left|\frac{1}{w}\right| f_{XY}\left(w, \frac{z}{w}\right)$$

- We wanted density of $Z = XY$.
- We used: $Z = XY$ and $W = Y$.
- We could have used: $Z = XY$ and $W = X$.
- This is invertible: $X = W$ and $Y = Z/W$.

$$J = \begin{vmatrix} 0 & 1 \\ \frac{1}{w} & \frac{-z}{w^2} \end{vmatrix} = -\frac{1}{w}$$

- This gives

$$\begin{aligned} f_{ZW}(z,w) &= \left|\frac{1}{w}\right| f_{XY}\left(w, \frac{z}{w}\right) \\ f_Z(z) &= \int_{-\infty}^{\infty} \left|\frac{1}{w}\right| f_{XY}\left(w, \frac{z}{w}\right) \, dw \end{aligned}$$

- We wanted density of $Z = XY$.
- We used: $Z = XY$ and $W = Y$.
- We could have used: $Z = XY$ and $W = X$.
- This is invertible: $X = W$ and $Y = Z/W$.

$$J = \begin{vmatrix} 0 & 1 \\ \frac{1}{w} & \frac{-z}{w^2} \end{vmatrix} = -\frac{1}{w}$$

- This gives

$$\begin{aligned} f_{ZW}(z, w) &= \left|\frac{1}{w}\right| f_{XY}\left(w, \frac{z}{w}\right) \\ f_Z(z) &= \int_{-\infty}^{\infty} \left|\frac{1}{w}\right| f_{XY}\left(w, \frac{z}{w}\right) \, dw \end{aligned}$$

- The f_Z should be same in both cases.

Distributions of quotients

- X, Y have joint density and $Z = X/Y$.

Distributions of quotients

- X, Y have joint density and $Z = X/Y$.
- We can take: $Z = X/Y \quad W = Y$

Distributions of quotients

- X, Y have joint density and $Z = X/Y$.
- We can take: $Z = X/Y \quad W = Y$
- This is invertible: $X = ZW \quad Y = W$

Distributions of quotients

- X, Y have joint density and $Z = X/Y$.
- We can take: $Z = X/Y \quad W = Y$
- This is invertible: $X = ZW \quad Y = W$

$$J = \begin{vmatrix} w & z \\ 0 & 1 \end{vmatrix} = w$$

Distributions of quotients

- X, Y have joint density and $Z = X/Y$.
- We can take: $Z = X/Y \quad W = Y$
- This is invertible: $X = ZW \quad Y = W$

$$J = \begin{vmatrix} w & z \\ 0 & 1 \end{vmatrix} = w$$

- Hence we get

$$f_{ZW}(z, w) = |w| \; f_{XY}(zw, w)$$

Distributions of quotients

- X, Y have joint density and $Z = X/Y$.
- We can take: $Z = X/Y \quad W = Y$
- This is invertible: $X = ZW \quad Y = W$

$$J = \begin{vmatrix} w & z \\ 0 & 1 \end{vmatrix} = w$$

- Hence we get

$$f_{ZW}(z, w) = |w| \; f_{XY}(zw, w)$$

- Thus we get the density of quotient as

$$f_Z(z) = \int_{-\infty}^{\infty} |w| \; f_{XY}(zw, w) \, dw$$

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = X/Y$.

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = X/Y$.
  Note $Z \in (0, \infty)$

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = X/Y$.
  Note $Z \in (0, \infty)$

$$f_Z(z) = \int_{-\infty}^{\infty} |w| \ f_X(zw) \ f_Y(w) \ dw$$

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = X/Y$.
  Note $Z \in (0, \infty)$

$$f_Z(z) = \int_{-\infty}^{\infty} |w|\ f_X(zw)\ f_Y(w)\ dw$$

- We need $0 < w < 1$ and $0 < zw < 1 \ \Rightarrow w < 1/z$.

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = X/Y$.
  Note $Z \in (0, \infty)$

$$f_Z(z) = \int_{-\infty}^{\infty} |w|\ f_X(zw)\ f_Y(w)\ dw$$

- We need $0 < w < 1$ and $0 < zw < 1 \ \Rightarrow w < 1/z$.
- So, when $z \leq 1$, w goes from 0 to 1; when $z > 1$, w goes from 0 to $1/z$.

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = X/Y$.
  Note $Z \in (0, \infty)$

$$f_Z(z) = \int_{-\infty}^{\infty} |w| \ f_X(zw) \ f_Y(w) \ dw$$

- We need $0 < w < 1$ and $0 < zw < 1 \ \Rightarrow w < 1/z$.
- So, when $z \leq 1$, w goes from 0 to 1; when $z > 1$, w goes from 0 to $1/z$.
- Hence we get density as

$$f_Z(z) = \begin{cases} \int_0^1 w \ dw = \frac{1}{2}, & \text{if } 0 < z \leq 1 \\ & \end{cases}$$

example

- Let X, Y be iid $U(0,\ 1)$. Let $Z = X/Y$.
  Note $Z \in (0, \infty)$

$$f_Z(z) = \int_{-\infty}^{\infty} |w|\ f_X(zw)\ f_Y(w)\ dw$$

- We need $0 < w < 1$ and $0 < zw < 1\ \Rightarrow w < 1/z$.
- So, when $z \leq 1$, w goes from 0 to 1; when $z > 1$, w goes from 0 to $1/z$.
- Hence we get density as

$$f_Z(z) = \begin{cases} \int_0^1 w\ dw = \frac{1}{2}, & \text{if } 0 < z \leq 1 \\ \int_0^{1/z} w\ dw = \frac{1}{2z^2}, & 1 < z < \infty \end{cases}$$

- X, Y have joint density and $Z = X/Y$

$$f_Z(z) = \int_{-\infty}^{\infty} |w| \; f_{XY}(zw, w) \; dw$$

- X, Y have joint density and $Z = X/Y$

$$f_Z(z) = \int_{-\infty}^{\infty} |w| \; f_{XY}(zw, w) \, dw$$

- SupposeX, Y are discrete and $Z = X/Y$

- X, Y have joint density and $Z = X/Y$

$$f_Z(z) = \int_{-\infty}^{\infty} |w| \; f_{XY}(zw, w) \, dw$$

- SupposeX, Y are discrete and $Z = X/Y$

$$f_Z(z) \;\; = \;\; P[Z = z] = P[X/Y = z]$$

- X, Y have joint density and $Z = X/Y$

$$f_Z(z) = \int_{-\infty}^{\infty} |w| \; f_{XY}(zw, w) \, dw$$

- SupposeX, Y are discrete and $Z = X/Y$

$$\begin{aligned} f_Z(z) &= P[Z = z] = P[X/Y = z] \\ &= \sum_y P[X = yz, Y = y] \end{aligned}$$

- X, Y have joint density and $Z = X/Y$

$$f_Z(z) = \int_{-\infty}^{\infty} |w| \ f_{XY}(zw, w) \, dw$$

- SupposeX, Y are discrete and $Z = X/Y$

$$\begin{aligned} f_Z(z) &= P[Z = z] = P[X/Y = z] \\ &= \sum_y P[X = yz, Y = y] \\ &= \sum_y f_{XY}(yz, y) \end{aligned}$$

- We chose: $Z = X/Y$ and $W = Y$.

- We chose: $Z = X/Y$ and $W = Y$.
- We could have taken: $Z = X/Y$ and $W = X$

- We chose: $Z = X/Y$ and $W = Y$.
- We could have taken: $Z = X/Y$ and $W = X$
- The inverse is: $X = W$ and $Y = W/Z$

- We chose: $Z = X/Y$ and $W = Y$.
- We could have taken: $Z = X/Y$ and $W = X$
- The inverse is: $X = W$ and $Y = W/Z$

$$J = \begin{vmatrix} 0 & 1 \\ -\frac{w}{z^2} & \frac{1}{z} \end{vmatrix} = -\frac{w}{z^2}$$

- We chose: $Z = X/Y$ and $W = Y$.
- We could have taken: $Z = X/Y$ and $W = X$
- The inverse is: $X = W$ and $Y = W/Z$

$$J = \begin{vmatrix} 0 & 1 \\ -\frac{w}{z^2} & \frac{1}{z} \end{vmatrix} = -\frac{w}{z^2}$$

- Thus we get the density of quotient as

$$f_Z(z) \;=\; \int_{-\infty}^{\infty} \left|\frac{w}{z^2}\right| \, f_{XY}\left(w, \frac{w}{z}\right) \, dw$$

- We chose: $Z = X/Y$ and $W = Y$.
- We could have taken: $Z = X/Y$ and $W = X$
- The inverse is: $X = W$ and $Y = W/Z$

$$J = \left| \begin{array}{cc} 0 & 1 \\ -\frac{w}{z^2} & \frac{1}{z} \end{array} \right| = -\frac{w}{z^2}$$

- Thus we get the density of quotient as

$$f_Z(z) \;=\; \int_{-\infty}^{\infty} \left|\frac{w}{z^2}\right| \, f_{XY}\left(w, \frac{w}{z}\right) \, dw$$

$$\text{put} \;\; t = \frac{w}{z} \;\; \Rightarrow dt = \frac{dw}{z}, \;\; w = tz$$

- We chose: $Z = X/Y$ and $W = Y$.
- We could have taken: $Z = X/Y$ and $W = X$
- The inverse is: $X = W$ and $Y = W/Z$

$$J = \begin{vmatrix} 0 & 1 \\ -\frac{w}{z^2} & \frac{1}{z} \end{vmatrix} = -\frac{w}{z^2}$$

- Thus we get the density of quotient as

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} \left|\frac{w}{z^2}\right| \, f_{XY}\left(w, \frac{w}{z}\right) \, dw \\ & \quad \text{put} \ \ t = \frac{w}{z} \ \ \Rightarrow dt = \frac{dw}{z}, \ \ w = tz \\ &= \int_{-\infty}^{\infty} |t| f_{XY}(tz, t) \, dt \end{aligned}$$

- We chose: $Z = X/Y$ and $W = Y$.
- We could have taken: $Z = X/Y$ and $W = X$
- The inverse is: $X = W$ and $Y = W/Z$

$$J = \begin{vmatrix} 0 & 1 \\ -\frac{w}{z^2} & \frac{1}{z} \end{vmatrix} = -\frac{w}{z^2}$$

- Thus we get the density of quotient as

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} \left|\frac{w}{z^2}\right| \; f_{XY}\left(w, \frac{w}{z}\right) \; dw \\ & \qquad \text{put} \;\; t = \frac{w}{z} \quad \Rightarrow dt = \frac{dw}{z}, \quad w = tz \\ &= \int_{-\infty}^{\infty} |t| f_{XY}(tz, t) \; dt \end{aligned}$$

- We can show that the density of quotient is same in both these approches.

Exchangeable Random Variables

- $X_1, X_2, \cdots, X_n$ are said to be exchangeable if their joint distribution is same as that of any permutation of them.

Exchangeable Random Variables

- $X_1, X_2, \cdots, X_n$ are said to be exchangeable if their joint distribution is same as that of any permutation of them.
- let $(i_1, \cdots, i_n)$ be a permutation of $(1, 2, \cdots, n)$. Then joint df of $(X_{i_1}, \cdots, X_{i_n})$ should be same as that $(X_1, \cdots, X_n)$

Exchangeable Random Variables

- $X_1, X_2, \cdots, X_n$ are said to be exchangeable if their joint distribution is same as that of any permutation of them.
- let $(i_1, \cdots, i_n)$ be a permutation of $(1, 2, \cdots, n)$. Then joint df of $(X_{i_1}, \cdots, X_{i_n})$ should be same as that $(X_1, \cdots, X_n)$
- Take $n = 3$. Suppose $F_{X_1 X_2 X_3}(a, b, c) = g(a, b, c)$. If they are exchangeable, then

Exchangeable Random Variables

- $X_1, X_2, \cdots, X_n$ are said to be exchangeable if their joint distribution is same as that of any permutation of them.
- let $(i_1, \cdots, i_n)$ be a permutation of $(1, 2, \cdots, n)$. Then joint df of $(X_{i_1}, \cdots, X_{i_n})$ should be same as that $(X_1, \cdots, X_n)$
- Take $n = 3$. Suppose $F_{X_1 X_2 X_3}(a, b, c) = g(a, b, c)$. If they are exchangeable, then

$$F_{X_2 X_3 X_1}(a, b, c) \quad = \quad P[X_2 \leq a, X_3 \leq b, X_1 \leq c]$$

Exchangeable Random Variables

- $X_1, X_2, \cdots, X_n$ are said to be exchangeable if their joint distribution is same as that of any permutation of them.
- let $(i_1, \cdots, i_n)$ be a permutation of $(1, 2, \cdots, n)$. Then joint df of $(X_{i_1}, \cdots, X_{i_n})$ should be same as that $(X_1, \cdots, X_n)$
- Take $n = 3$. Suppose $F_{X_1 X_2 X_3}(a, b, c) = g(a, b, c)$. If they are exchangeable, then

$$\begin{aligned} F_{X_2 X_3 X_1}(a, b, c) &= P[X_2 \le a, X_3 \le b, X_1 \le c] \\ &= P[X_1 \le c, X_2 \le a, X_3 \le b] \end{aligned}$$

Exchangeable Random Variables

- $X_1, X_2, \cdots, X_n$ are said to be exchangeable if their joint distribution is same as that of any permutation of them.
- let $(i_1, \cdots, i_n)$ be a permutation of $(1, 2, \cdots, n)$. Then joint df of $(X_{i_1}, \cdots, X_{i_n})$ should be same as that $(X_1, \cdots, X_n)$
- Take $n = 3$. Suppose $F_{X_1 X_2 X_3}(a, b, c) = g(a, b, c)$. If they are exchangeable, then

$$\begin{aligned} F_{X_2 X_3 X_1}(a, b, c) &= P[X_2 \leq a, X_3 \leq b, X_1 \leq c] \\ &= P[X_1 \leq c, X_2 \leq a, X_3 \leq b] \\ &= g(c, a, b) \end{aligned}$$

Exchangeable Random Variables

- $X_1, X_2, \cdots, X_n$ are said to be exchangeable if their joint distribution is same as that of any permutation of them.
- let $(i_1, \cdots, i_n)$ be a permutation of $(1, 2, \cdots, n)$. Then joint df of $(X_{i_1}, \cdots, X_{i_n})$ should be same as that $(X_1, \cdots, X_n)$
- Take $n = 3$. Suppose $F_{X_1 X_2 X_3}(a, b, c) = g(a, b, c)$. If they are exchangeable, then

$$\begin{aligned} F_{X_2 X_3 X_1}(a, b, c) &= P[X_2 \leq a, X_3 \leq b, X_1 \leq c] \\ &= P[X_1 \leq c, X_2 \leq a, X_3 \leq b] \\ &= g(c, a, b) = g(a, b, c) \end{aligned}$$

Exchangeable Random Variables

- $X_1, X_2, \cdots, X_n$ are said to be exchangeable if their joint distribution is same as that of any permutation of them.
- let $(i_1, \cdots, i_n)$ be a permutation of $(1, 2, \cdots, n)$. Then joint df of $(X_{i_1}, \cdots, X_{i_n})$ should be same as that $(X_1, \cdots, X_n)$
- Take $n = 3$. Suppose $F_{X_1 X_2 X_3}(a, b, c) = g(a, b, c)$. If they are exchangeable, then

$$\begin{aligned} F_{X_2 X_3 X_1}(a, b, c) &= P[X_2 \leq a, X_3 \leq b, X_1 \leq c] \\ &= P[X_1 \leq c, X_2 \leq a, X_3 \leq b] \\ &= g(c, a, b) = g(a, b, c) \end{aligned}$$

- The df or density should be "symmetric" in its variables if the random variables are exchangeable.

- Consider the density of three random variables

$$f(x, y, z) = \frac{2}{3}(x + y + z), \quad 0 < x, y, z < 1$$

- Consider the density of three random variables

$$f(x, y, z) = \frac{2}{3}(x + y + z), \quad 0 < x, y, z < 1$$

- They are exchangeable (because $f(x, y, z) = f(y, x, z)$)

- Consider the density of three random variables

$$f(x, y, z) = \frac{2}{3}(x + y + z), \quad 0 < x, y, z < 1$$

- They are exchangeable (because $f(x, y, z) = f(y, x, z)$)
- If random variables are exchangeable then they are identically distributed.

- Consider the density of three random variables

$$f(x, y, z) = \frac{2}{3}(x + y + z), \quad 0 < x, y, z < 1$$

- They are exchangeable (because $f(x, y, z) = f(y, x, z)$)
- If random variables are exchangeable then they are identically distributed.
  $F_{XYZ}(a, \infty, \infty) = F_{XYZ}(\infty, \infty, a) \quad \Rightarrow \quad F_X(a) = F_Z(a)$

- Consider the density of three random variables

$$f(x,y,z) = \frac{2}{3}(x+y+z), \quad 0 < x,y,z < 1$$

- They are exchangeable (because $f(x,y,z) = f(y,x,z)$)
- If random variables are exchangeable then they are identically distributed.
  $F_{XYZ}(a,\infty,\infty) = F_{XYZ}(\infty,\infty,a) \quad \Rightarrow \quad F_X(a) = F_Z(a)$
- The above example shows that exchangeable random variables need not be independent. The joint density is not factorizable.

- Consider the density of three random variables

$$f(x, y, z) = \frac{2}{3}(x + y + z), \quad 0 < x, y, z < 1$$

- They are exchangeable (because $f(x, y, z) = f(y, x, z)$)
- If random variables are exchangeable then they are identically distributed.
$F_{XYZ}(a, \infty, \infty) = F_{XYZ}(\infty, \infty, a) \;\; \Rightarrow \;\; F_X(a) = F_Z(a)$
- The above example shows that exchangeable random variables need not be independent. The joint density is not factorizable.

$$\int_0^1 \int_0^1 \frac{2}{3}(x + y + z) \; dy \; dz = \frac{2(x + 1)}{3}$$

- Consider the density of three random variables

$$f(x,y,z) = \frac{2}{3}(x+y+z), \quad 0 < x,y,z < 1$$

- They are exchangeable (because $f(x,y,z) = f(y,x,z)$)
- If random variables are exchangeable then they are identically distributed.
  $F_{XYZ}(a,\infty,\infty) = F_{XYZ}(\infty,\infty,a) \;\; \Rightarrow \;\; F_X(a) = F_Z(a)$
- The above example shows that exchangeable random variables need not be independent. The joint density is not factorizable.

$$\int_0^1 \int_0^1 \frac{2}{3}(x+y+z)\; dy\; dz = \frac{2(x+1)}{3}$$

- So, the joint density is not the product of marginals

Expectation of functions of multiple rv

Expectation of functions of multiple rv

- **Theorem**: Let $Z = g(X_1, \cdots X_n) = g(\mathbf{X})$. Then

$$E[Z] = \int_{\Re^n} g(\mathbf{x}) \; dF_{\mathbf{X}}(\mathbf{x})$$

Expectation of functions of multiple rv

- **Theorem**: Let $Z = g(X_1, \cdots X_n) = g(\mathbf{X})$. Then

$$E[Z] = \int_{\Re^n} g(\mathbf{x}) \; dF_{\mathbf{X}}(\mathbf{x})$$

- That is, if they have a joint density, then

$$E[Z] = \int_{\Re^n} g(\mathbf{x}) \; f_{\mathbf{X}}(\mathbf{x}) \; d\mathbf{x}$$

Expectation of functions of multiple rv

- **Theorem**: Let $Z = g(X_1, \cdots X_n) = g(\mathbf{X})$. Then

$$E[Z] = \int_{\Re^n} g(\mathbf{x}) \; dF_{\mathbf{X}}(\mathbf{x})$$

- That is, if they have a joint density, then

$$E[Z] = \int_{\Re^n} g(\mathbf{x}) \; f_{\mathbf{X}}(\mathbf{x}) \; d\mathbf{x}$$

- Similarly, if all X_i are discrete

$$E[Z] = \sum_{\mathbf{x}} g(\mathbf{x}) \; f_{\mathbf{X}}(\mathbf{x})$$

- Let $Z = X + Y$. Let X, Y have joint density f_{XY}

- Let $Z = X + Y$. Let X, Y have joint density f_{XY}

$$E[X+Y] \;=\; \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\; f_{XY}(x,y)\; dx\; dy$$

- Let $Z = X + Y$. Let X, Y have joint density f_{XY}

$$
\begin{aligned}
E[X+Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\; f_{XY}(x,y)\; dx\; dy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} \; f_{XY}(x,y)\; dy\; dx \\
&\qquad + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} \; f_{XY}(x,y)\; dx\; dy
\end{aligned}
$$

- Let $Z = X + Y$. Let X, Y have joint density f_{XY}

$$
\begin{aligned}
E[X+Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y) \, f_{XY}(x,y) \, dx \, dy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{XY}(x,y) \, dy \, dx \\
&\qquad + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{XY}(x,y) \, dx \, dy \\
&= \int_{-\infty}^{\infty} x \, f_X(x) \, dx + \int_{-\infty}^{\infty} y \, f_Y(y) \, dy
\end{aligned}
$$

- Let $Z = X + Y$. Let X, Y have joint density f_{XY}

$$
\begin{aligned}
E[X+Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\ f_{XY}(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty}\ f_{XY}(x,y)\ dy\ dx \\
&\qquad + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty}\ f_{XY}(x,y)\ dx\ dy \\
&= \int_{-\infty}^{\infty} x\ f_X(x)\ dx + \int_{-\infty}^{\infty} y\ f_Y(y)\ dy \\
&= E[X] + E[Y]
\end{aligned}
$$

- Let $Z = X + Y$. Let X, Y have joint density f_{XY}

$$\begin{aligned}
E[X+Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y) \, f_{XY}(x,y) \, dx \, dy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{XY}(x,y) \, dy \, dx \\
&\quad + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{XY}(x,y) \, dx \, dy \\
&= \int_{-\infty}^{\infty} x \, f_X(x) \, dx + \int_{-\infty}^{\infty} y \, f_Y(y) \, dy \\
&= E[X] + E[Y]
\end{aligned}$$

- Expectation is a linear operator.

- Let $Z = X + Y$. Let X, Y have joint density f_{XY}

$$
\begin{aligned}
E[X+Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y)\; f_{XY}(x,y)\; dx\; dy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} \; f_{XY}(x,y)\; dy\; dx \\
&\quad + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} \; f_{XY}(x,y)\; dx\; dy \\
&= \int_{-\infty}^{\infty} x\; f_X(x)\; dx + \int_{-\infty}^{\infty} y\; f_Y(y)\; dy \\
&= E[X] + E[Y]
\end{aligned}
$$

- Expectation is a linear operator.
- This is true for all random variables.

- We saw $E[X+Y] = E[X] + E[Y]$.

- We saw $E[X+Y] = E[X] + E[Y]$.
- Let us calculate $\text{Var}(X+Y)$.

- We saw $E[X+Y] = E[X] + E[Y]$.
- Let us calculate $\text{Var}(X+Y)$.

$$\text{Var}(X+Y) = E\left[((X+Y) - E[X+Y])^2\right]$$

- We saw $E[X+Y] = E[X] + E[Y]$.
- Let us calculate $\text{Var}(X+Y)$.

$$\begin{aligned}\text{Var}(X+Y) &= E\left[((X+Y) - E[X+Y])^2\right] \\ &= E\left[((X-EX) + (Y-EY))^2\right]\end{aligned}$$

- We saw $E[X+Y] = E[X] + E[Y]$.
- Let us calculate $\text{Var}(X+Y)$.

$$\begin{aligned}
\text{Var}(X+Y) &= E\left[((X+Y) - E[X+Y])^2\right] \\
&= E\left[((X-EX) + (Y-EY))^2\right] \\
&= E\left[(X-EX)^2\right] + E\left[(Y-EY)^2\right] \\
&\qquad +2E\left[(X-EX)(Y-EY)\right]
\end{aligned}$$

- We saw $E[X+Y] = E[X] + E[Y]$.
- Let us calculate $\text{Var}(X+Y)$.

$$
\begin{aligned}
\text{Var}(X+Y) &= E\left[((X+Y) - E[X+Y])^2\right] \\
&= E\left[((X - EX) + (Y - EY))^2\right] \\
&= E\left[(X - EX)^2\right] + E\left[(Y - EY)^2\right] \\
&\qquad +2E\left[(X - EX)(Y - EY)\right] \\
&= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)
\end{aligned}
$$

- We saw $E[X+Y] = E[X] + E[Y]$.
- Let us calculate $\text{Var}(X+Y)$.

$$\begin{aligned}
\text{Var}(X+Y) &= E\left[((X+Y) - E[X+Y])^2\right] \\
&= E\left[((X - EX) + (Y - EY))^2\right] \\
&= E\left[(X - EX)^2\right] + E\left[(Y - EY)^2\right] \\
&\qquad +2E\left[(X - EX)(Y - EY)\right] \\
&= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)
\end{aligned}$$

where we define **covariance** between X, Y as

$$\text{Cov}(X, Y) = E\left[(X - EX)(Y - EY)\right]$$

- We define **covariance** between X and Y by

$$\text{Cov}(X,Y) = E\left[(X - EX)(Y - EY)\right]$$

- We define **covariance** between X and Y by

$$\begin{aligned}\text{Cov}(X,Y) &= E\left[(X-EX)(Y-EY)\right] \\ &= E\left[XY - X(EY) - Y(EX) + EX\ EY\right]\end{aligned}$$

- We define **covariance** between X and Y by

$$\begin{aligned}
\text{Cov}(X,Y) &= E\left[(X-EX)(Y-EY)\right] \\
&= E\left[XY - X(EY) - Y(EX) + EX\ EY\right] \\
&= E[XY] - EX\ EY
\end{aligned}$$

- We define **covariance** between X and Y by

$$\begin{aligned}\text{Cov}(X,Y) &= E\left[(X-EX)(Y-EY)\right] \\ &= E\left[XY - X(EY) - Y(EX) + EX\ EY\right] \\ &= E[XY] - EX\ EY\end{aligned}$$

- Note that $\text{Cov}(X,Y)$ can be positive or negative

- We define **covariance** between X and Y by

$$\begin{aligned}
\text{Cov}(X,Y) &= E\left[(X-EX)(Y-EY)\right] \\
&= E\left[XY - X(EY) - Y(EX) + EX\ EY\right] \\
&= E[XY] - EX\ EY
\end{aligned}$$

- Note that $\text{Cov}(X,Y)$ can be positive or negative
- X and Y are said to be uncorrelated if $\text{Cov}(X,Y) = 0$

- We define **covariance** between X and Y by

$$\begin{aligned}\text{Cov}(X,Y) &= E\left[(X-EX)(Y-EY)\right] \\ &= E\left[XY - X(EY) - Y(EX) + EX\ EY\right] \\ &= E[XY] - EX\ EY\end{aligned}$$

- Note that $\text{Cov}(X,Y)$ can be positive or negative
- X and Y are said to be uncorrelated if $\text{Cov}(X,Y) = 0$
- If X and Y are uncorrelated then

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$$

- We define **covariance** between X and Y by

$$\begin{aligned}\text{Cov}(X,Y) &= E\left[(X-EX)(Y-EY)\right] \\ &= E\left[XY - X(EY) - Y(EX) + EX\ EY\right] \\ &= E[XY] - EX\ EY\end{aligned}$$

- Note that $\text{Cov}(X,Y)$ can be positive or negative
- X and Y are said to be uncorrelated if $\text{Cov}(X,Y) = 0$
- If X and Y are uncorrelated then

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$$

- Note that $E[X+Y] = E[X] + E[Y]$ for all random variables.

Example

- Consider the joint density

$$f_{XY}(x, y) = 2, \quad 0 < x < y < 1$$

Example

- Consider the joint density

$$f_{XY}(x, y) = 2, \;\; 0 < x < y < 1$$

- We want to calculate $\text{Cov}(X, Y)$

Example

- Consider the joint density

$$f_{XY}(x, y) = 2, \;\; 0 < x < y < 1$$

- We want to calculate $\text{Cov}(X, Y)$

$$EX = \int_0^1 \int_x^1 x \; 2 \; dy \; dx$$

Example

- Consider the joint density

$$f_{XY}(x, y) = 2, \;\; 0 < x < y < 1$$

- We want to calculate $\text{Cov}(X, Y)$

$$EX = \int_0^1 \int_x^1 x \; 2 \; dy \; dx = 2 \int_0^1 x \; (1 - x) \; dx = \frac{1}{3}$$

Example

- Consider the joint density

$$f_{XY}(x,y) = 2, \;\; 0 < x < y < 1$$

- We want to calculate Cov(X,Y)

$$EX = \int_0^1 \int_x^1 x \; 2 \; dy \; dx = 2 \int_0^1 x \; (1-x) \; dx = \frac{1}{3}$$

$$EY = \int_0^1 \int_0^y y \; 2 \; dx \; dy$$

Example

- Consider the joint density

$$f_{XY}(x,y) = 2, \;\; 0 < x < y < 1$$

- We want to calculate Cov(X, Y)

$$EX = \int_0^1 \int_x^1 x \; 2 \; dy \; dx = 2 \int_0^1 x \; (1 - x) \; dx = \frac{1}{3}$$

$$EY = \int_0^1 \int_0^y y \; 2 \; dx \; dy = 2 \int_0^1 y^2 \; dy = \frac{2}{3}$$

Example

- Consider the joint density

$$f_{XY}(x,y) = 2, \;\; 0 < x < y < 1$$

- We want to calculate Cov(X,Y)

$$EX = \int_0^1 \int_x^1 x \; 2 \; dy \; dx = 2 \int_0^1 x \; (1-x) \; dx = \frac{1}{3}$$

$$EY = \int_0^1 \int_0^y y \; 2 \; dx \; dy = 2 \int_0^1 y^2 \; dy = \frac{2}{3}$$

$$E[XY] = \int_0^1 \int_0^y xy \; 2 \; dx \; dy$$

Example

- Consider the joint density

$$f_{XY}(x,y) = 2, \;\; 0 < x < y < 1$$

- We want to calculate $\text{Cov}(X, Y)$

$$EX = \int_0^1 \int_x^1 x \; 2 \; dy \; dx = 2 \int_0^1 x \; (1-x) \; dx = \frac{1}{3}$$

$$EY = \int_0^1 \int_0^y y \; 2 \; dx \; dy = 2 \int_0^1 y^2 \; dy = \frac{2}{3}$$

$$E[XY] = \int_0^1 \int_0^y xy \; 2 \; dx \; dy = 2 \int_0^1 y \; \frac{y^2}{2} \; dy = \frac{1}{4}$$

Example

- Consider the joint density

$$f_{XY}(x, y) = 2, \ \ 0 < x < y < 1$$

- We want to calculate Cov(X, Y)

$$EX = \int_0^1 \int_x^1 x \; 2 \; dy \; dx = 2 \int_0^1 x \; (1 - x) \; dx = \frac{1}{3}$$

$$EY = \int_0^1 \int_0^y y \; 2 \; dx \; dy = 2 \int_0^1 y^2 \; dy = \frac{2}{3}$$

$$E[XY] = \int_0^1 \int_0^y xy \; 2 \; dx \; dy = 2 \int_0^1 y \; \frac{y^2}{2} \; dy = \frac{1}{4}$$

- Hence, Cov$(X, Y) = E[XY] - EX \; EY = \frac{1}{4} - \frac{2}{9} = \frac{1}{36}$

Independent random variables are uncorrelated

- Suppose X, Y are independent. Then

Independent random variables are uncorrelated

- Suppose X, Y are independent. Then

$$E[XY] = \int \int x \, y \, f_{XY}(x, y) \, dx \, dy$$

Independent random variables are uncorrelated

- Suppose X, Y are independent. Then

$$
\begin{aligned}
E[XY] &= \int \int x \; y \; f_{XY}(x,y) \; dx \; dy \\
&= \int \int x \; y \; f_X(x) \; f_Y(y) \; dx \; dy
\end{aligned}
$$

Independent random variables are uncorrelated

- Suppose X, Y are independent. Then

$$
\begin{aligned}
E[XY] &= \int\int x\ y\ f_{XY}(x,y)\ dx\ dy \\
&= \int\int x\ y\ f_X(x)\ f_Y(y)\ dx\ dy \\
&= \int x f_X(x)\ dx \int y f_Y(y)\ dy
\end{aligned}
$$

Independent random variables are uncorrelated

- Suppose X, Y are independent. Then

$$
\begin{aligned}
E[XY] &= \int\int x\ y\ f_{XY}(x,y)\ dx\ dy \\
&= \int\int x\ y\ f_X(x)\ f_Y(y)\ dx\ dy \\
&= \int x f_X(x)\ dx \int y f_Y(y)\ dy = EX\ EY
\end{aligned}
$$

Independent random variables are uncorrelated

- Suppose X, Y are independent. Then

$$
\begin{aligned}
E[XY] &= \int\int x\ y\ f_{XY}(x,y)\ dx\ dy \\
&= \int\int x\ y\ f_X(x)\ f_Y(y)\ dx\ dy \\
&= \int x f_X(x)\ dx \int y f_Y(y)\ dy = EX\ EY
\end{aligned}
$$

- Then, $\mathrm{Cov}(X, Y) = E[XY] - EX\ EY = 0$.

Independent random variables are uncorrelated

- Suppose X, Y are independent. Then

$$
\begin{aligned}
E[XY] &= \int \int x \; y \; f_{XY}(x,y) \; dx \; dy \\
&= \int \int x \; y \; f_X(x) \; f_Y(y) \; dx \; dy \\
&= \int x f_X(x) \; dx \int y f_Y(y) \; dy = EX \; EY
\end{aligned}
$$

- Then, $\text{Cov}(X,Y) = E[XY] - EX \; EY = 0$.
- X, Y independent $\;\Rightarrow\;$ X, Y uncorrelated

Uncorrelated random variables may not be independent

- Suppose $X \sim \mathcal{N}(0, 1)$ Then, $EX = EX^3 = 0$

Uncorrelated random variables may not be independent

- Suppose $X \sim \mathcal{N}(0,1)$ Then, $EX = EX^3 = 0$
- Let $Y = X^2$ Then,

$$E[XY] = EX^3 = 0$$

Uncorrelated random variables may not be independent

- Suppose $X \sim \mathcal{N}(0,1)$ Then, $EX = EX^3 = 0$
- Let $Y = X^2$ Then,

$$E[XY] = EX^3 = 0 = EX\ EY$$

Uncorrelated random variables may not be independent

- Suppose $X \sim \mathcal{N}(0,1)$ Then, $EX = EX^3 = 0$
- Let $Y = X^2$ Then,

$$E[XY] = EX^3 = 0 = EX\ EY$$

- Thus X, Y are uncorrelated.

Uncorrelated random variables may not be independent

- Suppose $X \sim \mathcal{N}(0,1)$ Then, $EX = EX^3 = 0$
- Let $Y = X^2$ Then,

$$E[XY] = EX^3 = 0 = EX\ EY$$

- Thus X, Y are uncorrelated.
- Are they independent?

Uncorrelated random variables may not be independent

- Suppose $X \sim \mathcal{N}(0,1)$ Then, $EX = EX^3 = 0$
- Let $Y = X^2$ Then,

$$E[XY] = EX^3 = 0 = EX\ EY$$

- Thus X, Y are uncorrelated.
- Are they independent? No

Uncorrelated random variables may not be independent

- Suppose $X \sim \mathcal{N}(0,1)$ Then, $EX = EX^3 = 0$
- Let $Y = X^2$ Then,

$$E[XY] = EX^3 = 0 = EX\ EY$$

- Thus X, Y are uncorrelated.
- Are they independent? No

  e.g.,

$$P[X > 2\,|Y < 1] = 0$$

Uncorrelated random variables may not be independent

- Suppose $X \sim \mathcal{N}(0,1)$ Then, $EX = EX^3 = 0$
- Let $Y = X^2$ Then,

$$E[XY] = EX^3 = 0 = EX\ EY$$

- Thus X, Y are uncorrelated.
- Are they independent? No
  e.g.,
  $$P[X > 2\ |Y < 1] = 0 \neq P[X > 2]$$

Uncorrelated random variables may not be independent

- Suppose $X \sim \mathcal{N}(0, 1)$ Then, $EX = EX^3 = 0$
- Let $Y = X^2$ Then,

$$E[XY] = EX^3 = 0 = EX\ EY$$

- Thus X, Y are uncorrelated.
- Are they independent? No
  e.g.,

$$P[X > 2 \mid Y < 1] = 0 \neq P[X > 2]$$

- X, Y are uncorrealted does not imply they are independent.

- We define the **correlation coefficient** of X, Y by

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\ \text{Var}(Y)}}$$

- We define the **correlation coefficient** of X, Y by

$$\rho_{XY} = \frac{\mathsf{Cov}(X, Y)}{\sqrt{\mathsf{Var}(X)\ \mathsf{Var}(Y)}}$$

- If X, Y are uncorrelated then $\rho_{XY} = 0$.

- We define the **correlation coefficient** of X, Y by

$$\rho_{XY} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\ \text{Var}(Y)}}$$

- If X, Y are uncorrelated then $\rho_{XY} = 0$.
- We will show that $|\rho_{XY}| \leq 1$

- We define the **correlation coefficient** of X, Y by

$$\rho_{XY} = \frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)\ \mathsf{Var}(Y)}}$$

- If X, Y are uncorrelated then $\rho_{XY} = 0$.
- We will show that $|\rho_{XY}| \leq 1$
- Hence $-1 \leq \rho_{XY} \leq 1,\ \forall X, Y$

- We have $E\left[(\alpha X + \beta Y)^2\right] \geq 0, \ \forall \alpha, \beta \in \Re$

- We have $E\left[(\alpha X+\beta Y)^2\right] \geq 0, \ \forall \alpha, \beta \in \Re$

$$\alpha^2 E[X^2] + \beta^2 E[Y^2] + 2\alpha\beta E[XY] \quad \geq 0, \quad \forall \alpha, \beta \in \Re$$

- We have $E\left[(\alpha X + \beta Y)^2\right] \geq 0, \ \forall \alpha, \beta \in \Re$

$$\alpha^2 E[X^2] + \beta^2 E[Y^2] + 2\alpha\beta E[XY] \quad \geq 0, \quad \forall \alpha, \beta \in \Re$$

$$\text{Take} \quad \alpha = -\frac{E[XY]}{E[X^2]}$$

- We have $E\left[(\alpha X+\beta Y)^2\right] \geq 0,\ \forall \alpha, \beta \in \Re$

$$\alpha^2 E[X^2] + \beta^2 E[Y^2] + 2\alpha\beta E[XY] \quad \geq 0, \quad \forall \alpha, \beta \in \Re$$

$$\text{Take} \quad \alpha = -\frac{E[XY]}{E[X^2]}$$

$$\frac{(E[XY])^2}{E[X^2]} +$$

- We have $E\left[(\alpha X+\beta Y)^2\right] \geq 0,\ \forall \alpha, \beta \in \Re$

$$\alpha^2 E[X^2] + \beta^2 E[Y^2] + 2\alpha\beta E[XY] \quad \geq 0, \quad \forall \alpha, \beta \in \Re$$

$$\text{Take} \quad \alpha = -\frac{E[XY]}{E[X^2]}$$

$$\frac{(E[XY])^2}{E[X^2]} + \beta^2 E[Y^2]$$

- We have $E\left[(\alpha X + \beta Y)^2\right] \geq 0, \ \forall \alpha, \beta \in \Re$

$$\alpha^2 E[X^2] + \beta^2 E[Y^2] + 2\alpha\beta E[XY] \quad \geq 0, \quad \forall \alpha, \beta \in \Re$$

$$\text{Take} \quad \alpha = -\frac{E[XY]}{E[X^2]}$$

$$\frac{(E[XY])^2}{E[X^2]} + \beta^2 E[Y^2] - 2\beta\frac{(E[XY])^2}{E[X^2]} \quad \geq 0, \quad \forall \beta \in \Re$$

- We have $E\left[(\alpha X+\beta Y)^2\right] \geq 0,\ \forall \alpha, \beta \in \Re$

$$\alpha^2 E[X^2] + \beta^2 E[Y^2] + 2\alpha\beta E[XY] \quad \geq 0, \quad \forall \alpha, \beta \in \Re$$

$$\text{Take} \quad \alpha = -\frac{E[XY]}{E[X^2]}$$

$$\frac{(E[XY])^2}{E[X^2]} + \beta^2 E[Y^2] - 2\beta \frac{(E[XY])^2}{E[X^2]} \quad \geq 0, \quad \forall \beta \in \Re$$

$$a\beta^2 + b\beta + c \geq 0,\ \forall \beta \ \Rightarrow b^2 - 4ac \leq 0$$

- We have $E\left[(\alpha X+\beta Y)^2\right] \geq 0, \ \forall \alpha, \beta \in \Re$

$$\alpha^2 E[X^2] + \beta^2 E[Y^2] + 2\alpha\beta E[XY] \geq 0, \quad \forall \alpha, \beta \in \Re$$

$$\text{Take} \quad \alpha = -\frac{E[XY]}{E[X^2]}$$

$$\frac{(E[XY])^2}{E[X^2]} + \beta^2 E[Y^2] - 2\beta \frac{(E[XY])^2}{E[X^2]} \geq 0, \quad \forall \beta \in \Re$$

$$\Rightarrow \quad 4\left(\frac{(E[XY])^2}{E[X^2]}\right)^2 - 4E[Y^2]\frac{(E[XY])^2}{E[X^2]} \leq 0$$

- We have $E\left[(\alpha X+\beta Y)^2\right] \geq 0,\ \forall \alpha, \beta \in \Re$

$$\alpha^2 E[X^2] + \beta^2 E[Y^2] + 2\alpha\beta E[XY] \geq 0, \quad \forall \alpha, \beta \in \Re$$

$$\text{Take} \quad \alpha = -\frac{E[XY]}{E[X^2]}$$

$$\frac{(E[XY])^2}{E[X^2]} + \beta^2 E[Y^2] - 2\beta \frac{(E[XY])^2}{E[X^2]} \geq 0, \quad \forall \beta \in \Re$$

$$\Rightarrow \quad 4\left(\frac{(E[XY])^2}{E[X^2]}\right)^2 - 4E[Y^2]\frac{(E[XY])^2}{E[X^2]} \leq 0$$

$$\Rightarrow \quad \left(\frac{(E[XY])^2}{E[X^2]}\right)^2 \leq \frac{E[Y^2](E[XY])^2}{E[X^2]}$$

- We have $E\left[(\alpha X+\beta Y)^2\right] \geq 0,\ \forall \alpha, \beta \in \Re$

$$\begin{aligned}
\alpha^2 E[X^2] + \beta^2 E[Y^2] + 2\alpha\beta E[XY] &\geq 0, \quad \forall \alpha, \beta \in \Re \\
\text{Take} \quad \alpha = -\frac{E[XY]}{E[X^2]} & \\
\frac{(E[XY])^2}{E[X^2]} + \beta^2 E[Y^2] - 2\beta \frac{(E[XY])^2}{E[X^2]} &\geq 0, \quad \forall \beta \in \Re
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad 4\left(\frac{(E[XY])^2}{E[X^2]}\right)^2 - 4E[Y^2]\frac{(E[XY])^2}{E[X^2]} &\leq 0 \\
\Rightarrow \quad \left(\frac{(E[XY])^2}{E[X^2]}\right)^2 \leq \frac{E[Y^2](E[XY])^2}{E[X^2]} & \\
\Rightarrow \quad \frac{(E[XY])^4}{(E[XY])^2} \leq \frac{E[Y^2](E[X^2])^2}{E[X^2]} &
\end{aligned}$$

- We have $E\left[(\alpha X + \beta Y)^2\right] \geq 0, \ \forall \alpha, \beta \in \Re$

$$\alpha^2 E[X^2] + \beta^2 E[Y^2] + 2\alpha\beta E[XY] \geq 0, \quad \forall \alpha, \beta \in \Re$$

$$\text{Take} \quad \alpha = -\frac{E[XY]}{E[X^2]}$$

$$\frac{(E[XY])^2}{E[X^2]} + \beta^2 E[Y^2] - 2\beta \frac{(E[XY])^2}{E[X^2]} \geq 0, \quad \forall \beta \in \Re$$

$$\Rightarrow \quad 4\left(\frac{(E[XY])^2}{E[X^2]}\right)^2 - 4E[Y^2]\frac{(E[XY])^2}{E[X^2]} \leq 0$$

$$\Rightarrow \left(\frac{(E[XY])^2}{E[X^2]}\right)^2 \leq \frac{E[Y^2](E[XY])^2}{E[X^2]}$$

$$\Rightarrow \frac{(E[XY])^4}{(E[XY])^2} \leq \frac{E[Y^2](E[X^2])^2}{E[X^2]}$$

$$\Rightarrow (E[XY])^2 \leq E[X^2]E[Y^2]$$

- We showed that

$$(E[XY])^2 \leq E[X^2]E[Y^2]$$

- We showed that
$$(E[XY])^2 \le E[X^2]E[Y^2]$$
- Take $X - EX$ in place of X and $Y - EY$ in place of Y in the above algebra.

- We showed that
$$(E[XY])^2 \le E[X^2]E[Y^2]$$
- Take $X - EX$ in place of X and $Y - EY$ in place of Y in the above algebra.
- This gives us
$$(E[(X - EX)(Y - EY)])^2 \le E[(X-EX)^2]E[(Y-EY)^2]$$

- We showed that

$$(E[XY])^2 \leq E[X^2]E[Y^2]$$

- Take $X - EX$ in place of X and $Y - EY$ in place of Y in the above algebra.
- This gives us

$$(E[(X - EX)(Y - EY)])^2 \leq E[(X-EX)^2]E[(Y-EY)^2]$$

$$\Rightarrow \quad (\text{Cov}(X, Y))^2 \leq \text{Var}(X)\text{Var}(Y)$$

- We showed that

$$(E[XY])^2 \leq E[X^2]E[Y^2]$$

- Take $X - EX$ in place of X and $Y - EY$ in place of Y in the above algebra.
- This gives us

$$(E[(X - EX)(Y - EY)])^2 \leq E[(X-EX)^2]E[(Y-EY)^2]$$

$$\Rightarrow \quad (\text{Cov}(X,Y))^2 \leq \text{Var}(X)\text{Var}(Y)$$

- Hence we get

$$\rho_{XY}^2 = \left(\frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}\right)^2 \leq 1$$

- We showed that

$$(E[XY])^2 \leq E[X^2]E[Y^2]$$

- Take $X - EX$ in place of X and $Y - EY$ in place of Y in the above algebra.
- This gives us

$$(E[(X - EX)(Y - EY)])^2 \leq E[(X-EX)^2]E[(Y-EY)^2]$$

$$\Rightarrow \quad (\mathsf{Cov}(X,Y))^2 \leq \mathsf{Var}(X)\mathsf{Var}(Y)$$

- Hence we get

$$\rho_{XY}^2 = \left(\frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)\mathsf{Var}(Y)}} \right)^2 \leq 1$$

- The equality holds here only if $E\left[(\alpha X + \beta Y)^2\right] = 0$

- We showed that

$$(E[XY])^2 \leq E[X^2]E[Y^2]$$

- Take $X - EX$ in place of X and $Y - EY$ in place of Y in the above algebra.
- This gives us

$$(E[(X - EX)(Y - EY)])^2 \leq E[(X - EX)^2]E[(Y - EY)^2]$$

$$\Rightarrow \quad (\text{Cov}(X, Y))^2 \leq \text{Var}(X)\text{Var}(Y)$$

- Hence we get

$$\rho_{XY}^2 = \left(\frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \right)^2 \leq 1$$

- The equality holds here only if $E\left[(\alpha X + \beta Y)^2\right] = 0$

Thus, $\quad |\rho_{XY}| = 1$ only if $\alpha X + \beta Y = 0$

- We showed that

$$(E[XY])^2 \leq E[X^2]E[Y^2]$$

- Take $X - EX$ in place of X and $Y - EY$ in place of Y in the above algebra.
- This gives us

$$(E[(X - EX)(Y - EY)])^2 \leq E[(X - EX)^2]E[(Y - EY)^2]$$

$$\Rightarrow \quad (\text{Cov}(X, Y))^2 \leq \text{Var}(X)\text{Var}(Y)$$

- Hence we get

$$\rho_{XY}^2 = \left(\frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}\right)^2 \leq 1$$

- The equality holds here only if $E\left[(\alpha X + \beta Y)^2\right] = 0$

Thus, $\quad |\rho_{XY}| = 1$ only if $\alpha X + \beta Y = 0$

- Correlation coefficient of X, Y is ± 1 only when Y is a linear function of X

Linear Least Squares Estimation

- Suppose we want to approximate Y as an affine function of X.

Linear Least Squares Estimation

- Suppose we want to approximate Y as an affine function of X.
- We want a, b to minimize $E\left[(Y - (aX + b))^2\right]$

Linear Least Squares Estimation

- Suppose we want to approximate Y as an affine function of X.
- We want a, b to minimize $E\left[(Y - (aX + b))^2\right]$
- For a fixed a, what is the b that minimizes $E\left[((Y - aX) - b)^2\right]$?

Linear Least Squares Estimation

- Suppose we want to approximate Y as an affine function of X.
- We want a, b to minimize $E\left[(Y - (aX + b))^2\right]$
- For a fixed a, what is the b that minimizes $E\left[((Y - aX) - b)^2\right]$?
- We know the best b here is: $b = E[Y - aX] = EY - aEX$.

Linear Least Squares Estimation

- Suppose we want to approximate Y as an affine function of X.
- We want a, b to minimize $E\left[(Y - (aX + b))^2\right]$
- For a fixed a, what is the b that minimizes $E\left[((Y - aX) - b)^2\right]$?
- We know the best b here is: $b = E[Y - aX] = EY - aEX$.
- So, we want to find the best a to minimize $J(a) = E\left[(Y - aX - (EY - aEX))^2\right]$

- We want to find a to minimize

$$J(a) \;=\; E\left[(Y - aX - (EY - aEX))^2\right]$$

- We want to find a to minimize

$$\begin{aligned} J(a) &= E\left[(Y - aX - (EY - aEX))^2\right] \\ &= E\left[((Y - EY) - a(X - EX))^2\right] \end{aligned}$$

- We want to find a to minimize

$$
\begin{aligned}
J(a) &= E\left[(Y - aX - (EY - aEX))^2\right] \\
&= E\left[((Y - EY) - a(X - EX))^2\right] \\
&= E\left[(Y - EY)^2 + a^2(X - EX)^2 - 2a(Y - EY)(X - EX)\right]
\end{aligned}
$$

- We want to find a to minimize

$$
\begin{aligned}
J(a) &= E\left[(Y - aX - (EY - aEX))^2\right] \\
&= E\left[((Y - EY) - a(X - EX))^2\right] \\
&= E\left[(Y - EY)^2 + a^2(X - EX)^2 - 2a(Y - EY)(X - EX)\right] \\
&= \mathsf{Var}(Y) + a^2\mathsf{Var}(X) - 2a\mathsf{Cov}(X, Y)
\end{aligned}
$$

- We want to find a to minimize

$$\begin{aligned} J(a) &= E\left[(Y - aX - (EY - aEX))^2\right] \\ &= E\left[((Y - EY) - a(X - EX))^2\right] \\ &= E\left[(Y - EY)^2 + a^2(X - EX)^2 - 2a(Y - EY)(X - EX)\right] \\ &= \text{Var}(Y) + a^2\text{Var}(X) - 2a\text{Cov}(X, Y) \end{aligned}$$

- So, the optimal a satisfies

$$2a\text{Var}(X) - 2\text{Cov}(X, Y) = 0$$

- We want to find a to minimize

$$\begin{aligned}
J(a) &= E\left[(Y - aX - (EY - aEX))^2\right] \\
&= E\left[((Y - EY) - a(X - EX))^2\right] \\
&= E\left[(Y - EY)^2 + a^2(X - EX)^2 - 2a(Y - EY)(X - EX)\right] \\
&= \text{Var}(Y) + a^2\text{Var}(X) - 2a\text{Cov}(X, Y)
\end{aligned}$$

- So, the optimal a satisfies

$$2a\text{Var}(X) - 2\text{Cov}(X, Y) = 0 \quad \Rightarrow \quad a = \frac{\text{Cov}(X, Y)}{\text{Var}(X)}$$

- The final mean square error, say, J^* is

$$J^* = \mathsf{Var}(Y) + a^2\mathsf{Var}(X) - 2a\mathsf{Cov}(X,Y)$$

- The final mean square error, say, J^* is

$$
\begin{aligned}
J^* &= \text{Var}(Y) + a^2\text{Var}(X) - 2a\text{Cov}(X, Y) \\
&= \text{Var}(Y) + \left(\frac{\text{Cov}(X, Y)}{\text{Var}(X)}\right)^2 \text{Var}(X) - 2\frac{\text{Cov}(X, Y)}{\text{Var}(X)}\ \text{Cov}(X, Y)
\end{aligned}
$$

- The final mean square error, say, J^* is

$$J^* = \text{Var}(Y) + a^2\text{Var}(X) - 2a\text{Cov}(X,Y)$$

$$= \text{Var}(Y) + \left(\frac{\text{Cov}(X,Y)}{\text{Var}(X)}\right)^2 \text{Var}(X) - 2\frac{\text{Cov}(X,Y)}{\text{Var}(X)}\,\text{Cov}(X,Y)$$

$$= \text{Var}(Y) - \frac{(\text{Cov}(X,Y))^2}{\text{Var}(X)}$$

- The final mean square error, say, J^* is

$$
\begin{aligned}
J^* &= \text{Var}(Y) + a^2\text{Var}(X) - 2a\text{Cov}(X,Y) \\
&= \text{Var}(Y) + \left(\frac{\text{Cov}(X,Y)}{\text{Var}(X)}\right)^2 \text{Var}(X) - 2\frac{\text{Cov}(X,Y)}{\text{Var}(X)}\ \text{Cov}(X,Y) \\
&= \text{Var}(Y) - \frac{(\text{Cov}(X,Y))^2}{\text{Var}(X)} \\
&= \text{Var}(Y)\ \left(1 - \frac{(\text{Cov}(X,Y))^2}{\text{Var}(Y)\ \text{Var}(X)}\right) \\
&= \text{Var}(Y)\left(1 - \rho_{XY}^2\right)
\end{aligned}
$$

- The best mean-square approximation of Y as a 'linear' function of X is

$$Y = \frac{\mathsf{Cov}(X,Y)}{\mathsf{Var}(X)}\, X \; + \; \left(EY - \frac{\mathsf{Cov}(X,Y)}{\mathsf{Var}(X)}\, EX \right)$$

- The best mean-square approximation of Y as a 'linear' function of X is

$$Y = \frac{\text{Cov}(X,Y)}{\text{Var}(X)} \, X \; + \; \left(EY - \frac{\text{Cov}(X,Y)}{\text{Var}(X)} \, EX \right)$$

- Called the line of regression of Y on X.

- The best mean-square approximation of Y as a 'linear' function of X is

$$Y = \frac{\text{Cov}(X,Y)}{\text{Var}(X)}\ X\ +\ \left(EY - \frac{\text{Cov}(X,Y)}{\text{Var}(X)}\ EX\right)$$

- Called the line of regression of Y on X.
- If $\text{cov}(X,Y) = 0$ then this reduces to approximating Y by a constant, EY.

- The best mean-square approximation of Y as a 'linear' function of X is

$$Y = \frac{\text{Cov}(X,Y)}{\text{Var}(X)}\, X \,+\, \left(EY - \frac{\text{Cov}(X,Y)}{\text{Var}(X)}\, EX \right)$$

- Called the line of regression of Y on X.
- If $\text{cov}(X,Y) = 0$ then this reduces to approximating Y by a constant, EY.
- The final mean square error is

$$\text{Var}(Y)\left(1 - \rho_{XY}^2\right)$$

- The best mean-square approximation of Y as a 'linear' function of X is

$$Y = \frac{\text{Cov}(X,Y)}{\text{Var}(X)}\, X \;+\; \left(EY - \frac{\text{Cov}(X,Y)}{\text{Var}(X)}\, EX\right)$$

- Called the line of regression of Y on X.
- If $\text{cov}(X,Y) = 0$ then this reduces to approximating Y by a constant, EY.
- The final mean square error is

$$\text{Var}(Y)\left(1 - \rho_{XY}^2\right)$$

- If $\rho_{XY} = \pm 1$ then the error is zero

- The best mean-square approximation of Y as a 'linear' function of X is

$$Y = \frac{\mathsf{Cov}(X,Y)}{\mathsf{Var}(X)}\, X \; + \; \left(EY - \frac{\mathsf{Cov}(X,Y)}{\mathsf{Var}(X)}\, EX \right)$$

- Called the line of regression of Y on X.
- If $\mathsf{cov}(X,Y) = 0$ then this reduces to approximating Y by a constant, EY.
- The final mean square error is

$$\mathsf{Var}(Y)\left(1 - \rho_{XY}^2\right)$$

- If $\rho_{XY} = \pm 1$ then the error is zero
- If $\rho_{XY} = 0$ the final error is $\mathsf{Var}(Y)$

- The covariance of X, Y is

$$\text{Cov}(X, Y) = E[(X - EX)\ (Y - EY)]$$

- The covariance of X, Y is

$$\mathrm{Cov}(X,Y) = E[(X - EX)\ (Y - EY)] = E[XY] - EX\ EY$$

- The covariance of X, Y is

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

  Note that $\text{Cov}(X, X) = \text{Var}(X)$

- The covariance of X, Y is

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

  Note that $\text{Cov}(X, X) = \text{Var}(X)$
- X, Y are called uncorrelated if $\text{Cov}(X, Y) = 0$.

- The covariance of X, Y is

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

  Note that $\text{Cov}(X, X) = \text{Var}(X)$
- X, Y are called uncorrelated if $\text{Cov}(X, Y) = 0$.
- X, Y independent $\Rightarrow$ X, Y uncorrelated.

- The covariance of X, Y is

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

  Note that $\text{Cov}(X,X) = \text{Var}(X)$
- X, Y are called uncorrelated if $\text{Cov}(X,Y) = 0$.
- X, Y independent $\Rightarrow$ X, Y uncorrelated.
- Uncorrelated random variables need not necessarily be independent

- The covariance of X, Y is

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

  Note that $\text{Cov}(X,X) = \text{Var}(X)$
- X, Y are called uncorrelated if $\text{Cov}(X,Y) = 0$.
- X, Y independent $\Rightarrow$ X, Y uncorrelated.
- Uncorrelated random variables need not necessarily be independent
- Covariance plays an important role in linear least squares estimation.

- The covariance of X, Y is

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

  Note that $\text{Cov}(X,X) = \text{Var}(X)$
- X, Y are called uncorrelated if $\text{Cov}(X,Y) = 0$.
- X, Y independent $\Rightarrow$ X, Y uncorrelated.
- Uncorrelated random variables need not necessarily be independent
- Covariance plays an important role in linear least squares estimation.
- Informally, covariance captures the 'linear dependence' between the two random variables.

Covariance Matrix

- Let $X_1, \cdots, X_n$ be random variables (on the same probability space)

Covariance Matrix

- Let $X_1, \cdots, X_n$ be random variables (on the same probability space)
- We represent them as a vector $\mathbf{X}$.

Covariance Matrix

- Let $X_1, \cdots, X_n$ be random variables (on the same probability space)
- We represent them as a vector $\mathbf{X}$.
- As a notation, all vectors are column vectors:

Covariance Matrix

- Let $X_1, \cdots, X_n$ be random variables (on the same probability space)
- We represent them as a vector $\mathbf{X}$.
- As a notation, all vectors are column vectors: $\mathbf{X} = (X_1, \cdots, X_n)^T$

Covariance Matrix

- Let $X_1, \cdots, X_n$ be random variables (on the same probability space)
- We represent them as a vector $\mathbf{X}$.
- As a notation, all vectors are column vectors: $\mathbf{X} = (X_1, \cdots, X_n)^T$
- We denote $E[\mathbf{X}] = (EX_1, \cdots, EX_n)^T$

Covariance Matrix

- Let $X_1, \cdots, X_n$ be random variables (on the same probability space)
- We represent them as a vector $\mathbf{X}$.
- As a notation, all vectors are column vectors: $\mathbf{X} = (X_1, \cdots, X_n)^T$
- We denote $E[\mathbf{X}] = (EX_1, \cdots, EX_n)^T$
- The $n \times n$ matrix whose $(i, j)^{th}$ element is Cov(X_i, X_j) is called the covariance matrix (or variance-covariance matrix) of $\mathbf{X}$. Denoted as $\Sigma_{\mathbf{X}}$ or Σ_X

Covariance Matrix

- Let $X_1, \cdots, X_n$ be random variables (on the same probability space)
- We represent them as a vector $\mathbf{X}$.
- As a notation, all vectors are column vectors: $\mathbf{X} = (X_1, \cdots, X_n)^T$
- We denote $E[\mathbf{X}] = (EX_1, \cdots, EX_n)^T$
- The $n \times n$ matrix whose $(i,j)^{th}$ element is $\text{Cov}(X_i, X_j)$ is called the covariance matrix (or variance-covariance matrix) of $\mathbf{X}$. Denoted as $\Sigma_{\mathbf{X}}$ or Σ_X

$$\Sigma_{\mathbf{X}} = \begin{bmatrix} \text{Cov}(X_1, X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & \text{Cov}(X_2, X_2) & \cdots & \text{Cov}(X_2, X_n) \\ \vdots & \vdots & \vdots & \vdots \\ \text{Cov}(X_n, X_1) & \text{Cov}(X_n, X_2) & \cdots & \text{Cov}(X_n, X_n) \end{bmatrix}$$

Covariance matrix

- If $\mathbf{a} = (a_1, \cdots, a_n)^T$ then
  $\mathbf{a}\ \mathbf{a}^T$ is a $n \times n$ matrix whose $(i,j)^{th}$ element is $a_i a_j$.

Covariance matrix

- If $\mathbf{a} = (a_1, \cdots, a_n)^T$ then
  $\mathbf{a}\ \mathbf{a}^T$ is a $n \times n$ matrix whose $(i,j)^{th}$ element is $a_i a_j$.
- Hence we get

$$\Sigma_{\mathbf{X}} = E\left[(\mathbf{X} - E\mathbf{X})\,(\mathbf{X} - E\mathbf{X})^T\right]$$

Covariance matrix

- If $\mathbf{a} = (a_1, \cdots, a_n)^T$ then
  $\mathbf{a}\,\mathbf{a}^T$ is a $n \times n$ matrix whose $(i,j)^{th}$ element is $a_i a_j$.
- Hence we get

$$\Sigma_{\mathbf{X}} = E\left[(\mathbf{X} - E\mathbf{X})\,(\mathbf{X} - E\mathbf{X})^T\right]$$

- This is because
  $\left((\mathbf{X} - E\mathbf{X})\,(\mathbf{X} - E\mathbf{X})^T\right)_{ij} = (X_i - EX_i)(X_j - EX_j)$

Covariance matrix

- If $\mathbf{a} = (a_1, \cdots, a_n)^T$ then
  $\mathbf{a}\,\mathbf{a}^T$ is a $n \times n$ matrix whose $(i,j)^{th}$ element is $a_i a_j$.
- Hence we get

$$\Sigma_{\mathbf{X}} = E\left[(\mathbf{X} - E\mathbf{X})\,(\mathbf{X} - E\mathbf{X})^T\right]$$

- This is because
  $\left((\mathbf{X} - E\mathbf{X})\,(\mathbf{X} - E\mathbf{X})^T\right)_{ij} = (X_i - EX_i)(X_j - EX_j)$
  and $(\Sigma_{\mathbf{X}})_{ij} = E[(X_i - EX_i)(X_j - EX_j)]$

- Recall the following about vectors and matrices

- Recall the following about vectors and matrices
- let $\mathbf{a}, \mathbf{b} \in \Re^n$ be column vectors. Then

$$\left(\mathbf{a}^T\mathbf{b}\right)^2 = \left(\mathbf{a}^T\mathbf{b}\right)^T \left(\mathbf{a}^T\mathbf{b}\right)$$

- Recall the following about vectors and matrices
- let $\mathbf{a}, \mathbf{b} \in \Re^n$ be column vectors. Then

$$\left(\mathbf{a}^T\mathbf{b}\right)^2 = \left(\mathbf{a}^T\mathbf{b}\right)^T \left(\mathbf{a}^T\mathbf{b}\right) = \mathbf{b}^T\mathbf{a}\,\mathbf{a}^T\mathbf{b}$$

- Recall the following about vectors and matrices
- let $\mathbf{a}, \mathbf{b} \in \Re^n$ be column vectors. Then

$$\left(\mathbf{a}^T\mathbf{b}\right)^2 = \left(\mathbf{a}^T\mathbf{b}\right)^T \left(\mathbf{a}^T\mathbf{b}\right) = \mathbf{b}^T\mathbf{a}\,\mathbf{a}^T\mathbf{b} = \mathbf{b}^T\left(\mathbf{a}\,\mathbf{a}^T\right)\mathbf{b}$$

- Recall the following about vectors and matrices
- let $\mathbf{a}, \mathbf{b} \in \Re^n$ be column vectors. Then

$$\left(\mathbf{a}^T\mathbf{b}\right)^2 = \left(\mathbf{a}^T\mathbf{b}\right)^T \left(\mathbf{a}^T\mathbf{b}\right) = \mathbf{b}^T\mathbf{a}\,\mathbf{a}^T\mathbf{b} = \mathbf{b}^T\left(\mathbf{a}\,\mathbf{a}^T\right)\mathbf{b}$$

- Let A be an $n \times n$ matrix with elements a_{ij}. Then

$$\mathbf{b}^T A\mathbf{b} = \sum_{i,j=1}^{n} b_i b_j a_{ij}$$

where $\mathbf{b} = (b_1, \cdots, b_n)^T$

- Recall the following about vectors and matrices
- let $\mathbf{a}, \mathbf{b} \in \Re^n$ be column vectors. Then

$$\left(\mathbf{a}^T\mathbf{b}\right)^2 = \left(\mathbf{a}^T\mathbf{b}\right)^T \left(\mathbf{a}^T\mathbf{b}\right) = \mathbf{b}^T\mathbf{a}\,\mathbf{a}^T\mathbf{b} = \mathbf{b}^T\left(\mathbf{a}\,\mathbf{a}^T\right)\mathbf{b}$$

- Let A be an $n \times n$ matrix with elements a_{ij}. Then

$$\mathbf{b}^T A\mathbf{b} = \sum_{i,j=1}^{n} b_i b_j a_{ij}$$

  where $\mathbf{b} = (b_1, \cdots, b_n)^T$
- A is said to be positive semidefinite if $\mathbf{b}^T A\mathbf{b} \geq 0, \; \forall \mathbf{b}$

- Σ_X is a real symmetric matrix

- Σ_X is a real symmetric matrix
- It is positive semidefinite.

- Σ_X is a real symmetric matrix
- It is positive semidefinite.
- Let $\mathbf{a} \in \Re^n$ and let $Y = \mathbf{a}^T \mathbf{X}$.

- Σ_X is a real symmetric matrix
- It is positive semidefinite.
- Let $\mathbf{a} \in \Re^n$ and let $Y = \mathbf{a}^T \mathbf{X}$.
- Then, $EY = \mathbf{a}^T E\mathbf{X}$. We get variance of Y as

$$\text{Var}(Y) \;=\; E[(Y - EY)^2] = E\left[\left(\mathbf{a}^T \mathbf{X} - \mathbf{a}^T E\mathbf{X}\right)^2\right]$$

- Σ_X is a real symmetric matrix
- It is positive semidefinite.
- Let $\mathbf{a} \in \Re^n$ and let $Y = \mathbf{a}^T \mathbf{X}$.
- Then, $EY = \mathbf{a}^T E\mathbf{X}$. We get variance of Y as

$$\begin{aligned} \mathrm{Var}(Y) &= E[(Y - EY)^2] = E\left[\left(\mathbf{a}^T\mathbf{X} - \mathbf{a}^T E\mathbf{X}\right)^2\right] \\ &= E\left[\left(\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\right)^2\right] \end{aligned}$$

- Σ_X is a real symmetric matrix
- It is positive semidefinite.
- Let $\mathbf{a} \in \Re^n$ and let $Y = \mathbf{a}^T\mathbf{X}$.
- Then, $EY = \mathbf{a}^T E\mathbf{X}$. We get variance of Y as

$$\begin{aligned}
\text{Var}(Y) &= E[(Y - EY)^2] = E\left[\left(\mathbf{a}^T\mathbf{X} - \mathbf{a}^T E\mathbf{X}\right)^2\right] \\
&= E\left[\left(\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\right)^2\right] \\
&= E\left[\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\,(\mathbf{X} - E\mathbf{X})^T\mathbf{a}\right]
\end{aligned}$$

- Σ_X is a real symmetric matrix
- It is positive semidefinite.
- Let $\mathbf{a} \in \Re^n$ and let $Y = \mathbf{a}^T\mathbf{X}$.
- Then, $EY = \mathbf{a}^T E\mathbf{X}$. We get variance of Y as

$$
\begin{aligned}
\mathrm{Var}(Y) &= E[(Y - EY)^2] = E\left[\left(\mathbf{a}^T\mathbf{X} - \mathbf{a}^T E\mathbf{X}\right)^2\right] \\
&= E\left[\left(\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\right)^2\right] \\
&= E\left[\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\left(\mathbf{X} - E\mathbf{X}\right)^T\mathbf{a}\right] \\
&= \mathbf{a}^T E\left[(\mathbf{X} - E\mathbf{X})\left(\mathbf{X} - E\mathbf{X}\right)^T\right]\mathbf{a}
\end{aligned}
$$

- Σ_X is a real symmetric matrix
- It is positive semidefinite.
- Let $\mathbf{a} \in \Re^n$ and let $Y = \mathbf{a}^T \mathbf{X}$.
- Then, $EY = \mathbf{a}^T E\mathbf{X}$. We get variance of Y as

$$
\begin{aligned}
\text{Var}(Y) &= E[(Y - EY)^2] = E\left[\left(\mathbf{a}^T\mathbf{X} - \mathbf{a}^T E\mathbf{X}\right)^2\right] \\
&= E\left[\left(\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\right)^2\right] \\
&= E\left[\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\left(\mathbf{X} - E\mathbf{X}\right)^T\mathbf{a}\right] \\
&= \mathbf{a}^T E\left[(\mathbf{X} - E\mathbf{X})\left(\mathbf{X} - E\mathbf{X}\right)^T\right]\mathbf{a} \\
&= \mathbf{a}^T\Sigma_X\mathbf{a}
\end{aligned}
$$

- Σ_X is a real symmetric matrix
- It is positive semidefinite.
- Let $\mathbf{a} \in \Re^n$ and let $Y = \mathbf{a}^T \mathbf{X}$.
- Then, $EY = \mathbf{a}^T E\mathbf{X}$. We get variance of Y as

$$
\begin{aligned}
\text{Var}(Y) &= E[(Y - EY)^2] = E\left[\left(\mathbf{a}^T\mathbf{X} - \mathbf{a}^T E\mathbf{X}\right)^2\right] \\
&= E\left[\left(\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\right)^2\right] \\
&= E\left[\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\left(\mathbf{X} - E\mathbf{X}\right)^T\mathbf{a}\right] \\
&= \mathbf{a}^T E\left[(\mathbf{X} - E\mathbf{X})\left(\mathbf{X} - E\mathbf{X}\right)^T\right]\mathbf{a} \\
&= \mathbf{a}^T \Sigma_X \mathbf{a}
\end{aligned}
$$

- This gives $\mathbf{a}^T \Sigma_X \mathbf{a} \geq 0, \ \forall \mathbf{a}$

- Σ_X is a real symmetric matrix
- It is positive semidefinite.
- Let $\mathbf{a} \in \Re^n$ and let $Y = \mathbf{a}^T\mathbf{X}$.
- Then, $EY = \mathbf{a}^T E\mathbf{X}$. We get variance of Y as

$$\begin{aligned}
\text{Var}(Y) &= E[(Y - EY)^2] = E\left[\left(\mathbf{a}^T\mathbf{X} - \mathbf{a}^T E\mathbf{X}\right)^2\right] \\
&= E\left[\left(\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\right)^2\right] \\
&= E\left[\mathbf{a}^T(\mathbf{X} - E\mathbf{X})\,(\mathbf{X} - E\mathbf{X})^T\mathbf{a}\right] \\
&= \mathbf{a}^T E\left[(\mathbf{X} - E\mathbf{X})\,(\mathbf{X} - E\mathbf{X})^T\right]\mathbf{a} \\
&= \mathbf{a}^T\Sigma_X\mathbf{a}
\end{aligned}$$

- This gives $\mathbf{a}^T\Sigma_X\mathbf{a} \geq 0, \ \forall\mathbf{a}$
- This shows Σ_X is positive semidefinite

- $Y = \mathbf{a}^T \mathbf{X} = \sum_i a_i X_i$

- $Y = \mathbf{a}^T\mathbf{X} = \sum_i a_i X_i$ – linear combination of X_i's.

- $Y = \mathbf{a}^T \mathbf{X} = \sum_i a_i X_i$ – linear combination of X_i's.
- We know how to find its mean and variance

- $Y = \mathbf{a}^T \mathbf{X} = \sum_i a_i X_i$ – linear combination of X_i's.
- We know how to find its mean and variance

$$EY = \mathbf{a}^T E\mathbf{X} = \sum_i a_i EX_i;$$

- $Y = \mathbf{a}^T \mathbf{X} = \sum_i a_i X_i$ – linear combination of X_i's.
- We know how to find its mean and variance

$$
\begin{aligned}
EY &= \mathbf{a}^T E\mathbf{X} = \sum_i a_i E X_i; \\
\text{Var}(Y) &= \mathbf{a}^T \Sigma_X \mathbf{a} = \sum_{i,j} a_i a_j \text{Cov}(X_i, X_j)
\end{aligned}
$$

- $Y = \mathbf{a}^T \mathbf{X} = \sum_i a_i X_i$ – linear combination of X_i's.
- We know how to find its mean and variance

$$\begin{aligned} EY &= \mathbf{a}^T E\mathbf{X} = \sum_i a_i E X_i; \\ \mathsf{Var}(Y) &= \mathbf{a}^T \Sigma_X \mathbf{a} = \sum_{i,j} a_i a_j \mathsf{Cov}(X_i, X_j) \end{aligned}$$

- Specifically, by taking all components of $\mathbf{a}$ to be 1, we get

$$\mathsf{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i,j=1}^{n} \mathsf{Cov}(X_i, X_j)$$

- $Y = \mathbf{a}^T \mathbf{X} = \sum_i a_i X_i$ – linear combination of X_i's.
- We know how to find its mean and variance

$$\begin{aligned} EY &= \mathbf{a}^T E\mathbf{X} = \sum_i a_i EX_i; \\ \mathsf{Var}(Y) &= \mathbf{a}^T \Sigma_X \mathbf{a} = \sum_{i,j} a_i a_j \mathsf{Cov}(X_i, X_j) \end{aligned}$$

- Specifically, by taking all components of $\mathbf{a}$ to be 1, we get

$$\mathsf{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i,j=1}^n \mathsf{Cov}(X_i, X_j) = \sum_{i=1}^n \mathsf{Var}(X_i) + \sum_{i=1}^n \sum_{j \neq i} \mathsf{Cov}(X_i, X_j)$$

- $Y = \mathbf{a}^T \mathbf{X} = \sum_i a_i X_i$ – linear combination of X_i's.
- We know how to find its mean and variance

$$
\begin{aligned}
EY &= \mathbf{a}^T E\mathbf{X} = \sum_i a_i EX_i; \\
\mathsf{Var}(Y) &= \mathbf{a}^T \Sigma_X \mathbf{a} = \sum_{i,j} a_i a_j \mathsf{Cov}(X_i, X_j)
\end{aligned}
$$

- Specifically, by taking all components of $\mathbf{a}$ to be 1, we get

$$
\mathsf{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i,j=1}^n \mathsf{Cov}(X_i, X_j) = \sum_{i=1}^n \mathsf{Var}(X_i) + \sum_{i=1}^n \sum_{j \neq i} \mathsf{Cov}(X_i, X_j)
$$

- If X_i are independent, variance of sum is sum of variances.

- Covariance matrix Σ_X positive semidefinite because

$$\mathbf{a}^T \Sigma_X \, \mathbf{a} = \mathsf{Var}(\mathbf{a}^T \mathbf{X}) \geq 0$$

- Covariance matrix Σ_X positive semidefinite because

$$\mathbf{a}^T \Sigma_X \ \mathbf{a} = \mathsf{Var}(\mathbf{a}^T \mathbf{X}) \geq 0$$

- Σ_X would be positive definite if $\mathbf{a}^T \Sigma_X \ \mathbf{a} > 0, \ \forall \mathbf{a} \neq 0$

- Covariance matrix Σ_X positive semidefinite because

$$\mathbf{a}^T \Sigma_X \, \mathbf{a} = \mathsf{Var}(\mathbf{a}^T \mathbf{X}) \geq 0$$

- Σ_X would be positive definite if $\mathbf{a}^T \Sigma_X \, \mathbf{a} > 0, \ \forall \mathbf{a} \neq 0$
- It would fail to be positive definite if $\mathsf{Var}(\mathbf{a}^T \mathbf{X}) = 0$ for some nonzero $\mathbf{a}$.

- Covariance matrix Σ_X positive semidefinite because

$$\mathbf{a}^T \Sigma_X \ \mathbf{a} = \mathsf{Var}(\mathbf{a}^T \mathbf{X}) \geq 0$$

- Σ_X would be positive definite if $\mathbf{a}^T \Sigma_X \ \mathbf{a} > 0, \ \forall \mathbf{a} \neq 0$
- It would fail to be positive definite if $\mathsf{Var}(\mathbf{a}^T \mathbf{X}) = 0$ for some nonzero $\mathbf{a}$.
- $\mathsf{Var}(Z) = E[(Z - EZ)^2] = 0$ implies $Z = EZ$, a constant.

- Covariance matrix Σ_X positive semidefinite because

$$\mathbf{a}^T \Sigma_X \, \mathbf{a} = \mathsf{Var}(\mathbf{a}^T \mathbf{X}) \geq 0$$

- Σ_X would be positive definite if $\mathbf{a}^T \Sigma_X \, \mathbf{a} > 0, \ \forall \mathbf{a} \neq 0$
- It would fail to be positive definite if $\mathsf{Var}(\mathbf{a}^T \mathbf{X}) = 0$ for some nonzero $\mathbf{a}$.
- $\mathsf{Var}(Z) = E[(Z - EZ)^2] = 0$ implies $Z = EZ$, a constant.
- Hence, Σ_X fails to be positive definite only if there is a non-zero linear combination of X_i's that is a constant.

- Covariance matrix is a real symmetric positive semidefinite matrix

- Covariance matrix is a real symmetric positive semidefinite matrix
- It have real and non-negative eigen values.

- Covariance matrix is a real symmetric positive semidefinite matrix
- It have real and non-negative eigen values.
- It would have n linearly independent eigen vectors.

- Covariance matrix is a real symmetric positive semidefinite matrix
- It have real and non-negative eigen values.
- It would have n linearly independent eigen vectors.
- These also have some interesting roles.

- Covariance matrix is a real symmetric positive semidefinite matrix
- It have real and non-negative eigen values.
- It would have n linearly independent eigen vectors.
- These also have some interesting roles.
- We consider one simple example.

- Let $Y = \mathbf{a}^T \mathbf{X}$ and assume $||\mathbf{a}|| = 1$

- Let $Y = \mathbf{a}^T \mathbf{X}$ and assume $||\mathbf{a}|| = 1$
- Y is projection of $\mathbf{X}$ along the direction $\mathbf{a}$.

- Let $Y = \mathbf{a}^T \mathbf{X}$ and assume $||\mathbf{a}|| = 1$
- Y is projection of $\mathbf{X}$ along the direction $\mathbf{a}$.
- Suppose we want to find a direction along which variance is maximized

- Let $Y = \mathbf{a}^T \mathbf{X}$ and assume $||\mathbf{a}|| = 1$
- Y is projection of $\mathbf{X}$ along the direction $\mathbf{a}$.
- Suppose we want to find a direction along which variance is maximized
- We want to maximize $\mathbf{a}^T \Sigma_X \, \mathbf{a}$ subject to $\mathbf{a}^T \mathbf{a} = 1$

- Let $Y = \mathbf{a}^T\mathbf{X}$ and assume $||\mathbf{a}|| = 1$
- Y is projection of $\mathbf{X}$ along the direction $\mathbf{a}$.
- Suppose we want to find a direction along which variance is maximized
- We want to maximize $\mathbf{a}^T\Sigma_X\ \mathbf{a}$ subject to $\mathbf{a}^T\mathbf{a} = 1$
- The lagrangian is $\mathbf{a}^T\Sigma_X\ \mathbf{a} + \eta(1 - \mathbf{a}^T\mathbf{a})$

- Let $Y = \mathbf{a}^T \mathbf{X}$ and assume $||\mathbf{a}|| = 1$
- Y is projection of $\mathbf{X}$ along the direction $\mathbf{a}$.
- Suppose we want to find a direction along which variance is maximized
- We want to maximize $\mathbf{a}^T \Sigma_X \, \mathbf{a}$ subject to $\mathbf{a}^T \mathbf{a} = 1$
- The lagrangian is $\mathbf{a}^T \Sigma_X \, \mathbf{a} + \eta(1 - \mathbf{a}^T \mathbf{a})$
- Equating the gradient to zero, we get

$$\Sigma_X \, \mathbf{a} = \eta \mathbf{a}$$

- Let $Y = \mathbf{a}^T \mathbf{X}$ and assume $||\mathbf{a}|| = 1$
- Y is projection of $\mathbf{X}$ along the direction $\mathbf{a}$.
- Suppose we want to find a direction along which variance is maximized
- We want to maximize $\mathbf{a}^T \Sigma_X \ \mathbf{a}$ subject to $\mathbf{a}^T \mathbf{a} = 1$
- The lagrangian is $\mathbf{a}^T \Sigma_X \ \mathbf{a} + \eta(1 - \mathbf{a}^T \mathbf{a})$
- Equating the gradient to zero, we get

$$\Sigma_X \ \mathbf{a} = \eta \mathbf{a}$$

- So, $\mathbf{a}$ should be an eigen vector (with eigen value η).

- Let $Y = \mathbf{a}^T \mathbf{X}$ and assume $||\mathbf{a}|| = 1$
- Y is projection of $\mathbf{X}$ along the direction $\mathbf{a}$.
- Suppose we want to find a direction along which variance is maximized
- We want to maximize $\mathbf{a}^T \Sigma_X \ \mathbf{a}$ subject to $\mathbf{a}^T \mathbf{a} = 1$
- The lagrangian is $\mathbf{a}^T \Sigma_X \ \mathbf{a} + \eta(1 - \mathbf{a}^T \mathbf{a})$
- Equating the gradient to zero, we get

$$\Sigma_X \ \mathbf{a} = \eta \mathbf{a}$$

- So, $\mathbf{a}$ should be an eigen vector (with eigen value η).
- Then the variance would be $\mathbf{a}^T \Sigma_X \ \mathbf{a} = \eta \mathbf{a}^T \mathbf{a} = \eta$

- Let $Y = \mathbf{a}^T \mathbf{X}$ and assume $||\mathbf{a}|| = 1$
- Y is projection of $\mathbf{X}$ along the direction $\mathbf{a}$.
- Suppose we want to find a direction along which variance is maximized
- We want to maximize $\mathbf{a}^T \Sigma_X \ \mathbf{a}$ subject to $\mathbf{a}^T \mathbf{a} = 1$
- The lagrangian is $\mathbf{a}^T \Sigma_X \ \mathbf{a} + \eta(1 - \mathbf{a}^T \mathbf{a})$
- Equating the gradient to zero, we get

$$\Sigma_X \ \mathbf{a} = \eta \mathbf{a}$$

- So, $\mathbf{a}$ should be an eigen vector (with eigen value η).
- Then the variance would be $\mathbf{a}^T \Sigma_X \ \mathbf{a} = \eta \mathbf{a}^T \mathbf{a} = \eta$
- Hence the direction is the eigen vector corresponding to the highest eigen value.