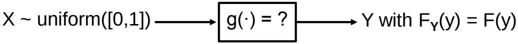
Lecture 7

- **Read:** Chapter 3.8, Chapter 4.1, 4.4-4.11.
- Continuous Random Variables (continued)
  - Probability Models of Derived Random Variables
  - Conditioning a Continuous Random Variable
- Multiple Continuous Random Variables
  - Joint Cumulative Distribution Function
  - Joint Probability Density Function
  - Marginal Probability Density Function
  - Functions of Two Random Variables
  - Expected Values
  - Conditioning by an Event/Conditioning by a Random Variable
  - Independent Random Variables
  - Jointly (Bivariate) Gaussian Random Variables

Application: Generating RVs on a Computer: Setup

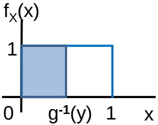
- Suppose your computer can generate $X \sim$ uniform[0,1] RVs (e.g., do a random() call).
- How do we generate some other random variable, say Y, with a given CDF, say $F(\cdot)$?

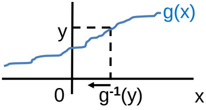
Application: Generating RVs on a Computer: Approach



Suppose that g() is increasing.

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





$P[X \le g^{-1}(y)] = g^{-1}(y)$ when $0 \le g^{-1}(y) \le 1$

$y = g(x)$

$g^{-1}(y) = g^{-1}(g(x)) = x$

- Our goal is to make $F_Y(y) = g^{-1}(y) = \underbrace{F_Y(y)}_{\text{prespecified CDF}}$.
- Thus, $g^{-1}(y) = F(y)$ and $\underbrace{g(g^{-1}(y))}_{y=g(F(y))} = g(F(y))$
- Thus, $g = F^{-1}$, the inverse of the specified CDF.

Application: Generating RVs on a Computer: Example

- How do we generate exponential RVs based on uniform RVs?

. . . .

- Recall Y is exponential(a) if

$$F_Y(y) = F(y) = \begin{cases} 1 - e^{-ay} & , y \geq 0 \\ 0 & , \text{otherwise} \end{cases}$$

- Since $g = F^{-1}$, if $x = 1 - e^{-ay}$,

$$\begin{aligned} x - 1 &= -e^{-ay} \\ 1 - x &= e^{-ay} \\ \ln(1 - x) &= -ay \\ \frac{\ln(1 - x)}{-a} &= y \end{aligned}$$

- So, $g(x) = \frac{\ln(1-x)}{-a}$.
- If $X \sim$ uniform[0,1], then $Y = \frac{\ln(1-X)}{-a} \sim$ exp(a).
- **Note:** $Y = \frac{\ln(X)}{-a}$ also works! (because if X is uniform on [0,1], then so is 1 - X)

Reminder: “Functions of Discrete RVs”

- Suppose X is a discrete RV with range S_X and PMF $p_X(x)$.
- Let $Y = g(X)$.
- Then, Y is also discrete with $S_Y = \{g(x)|x \in S_X\}$ and

$$p_Y(y) = \sum_{\substack{x:g(x)=y \\ x \in S_X}} p_X(x)$$

- **Example:** Suppose $X \sim$ uniform {-1,0,1,2} and $Y = X^2$. Then, $S_Y = \{0,1,4\}$ and

$$p_Y(0) = p_Y(4) = 1/4$$
$$p_Y(1) = 1/2$$

Summary: Possibilities for Derived Distributions

$$X \longrightarrow \boxed{g(X)} \longrightarrow Y$$

- discrete $\longrightarrow$ discrete
- continuous $\longrightarrow$ $\left\{\begin{array}{l}\text{continuous}\\ \text{discrete}\\ \text{mixed}\end{array}\right.$
- mixed $\longrightarrow$ $\left\{\begin{array}{l}\text{mixed}\\ \text{discrete}\end{array}\right.$

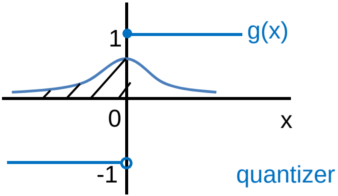
There is at least one value assumed with positive probability; cannot be continuous

Getting a Discrete RV from a Continuous RV

- **Example:** Let $X \sim N(0,1)$ and let

$$g(x) = \begin{cases} 1 & , x \geq 0 \\ -1 & , x < 0 \end{cases}$$

Let $Y = g(X)$. What is the PDF/PMF of Y?



Note S_Y = {-1, 1}
and P[Y = -1] = P[X < 0] = 1/2
P[Y = 1] = 1/2

- **Remark:** In general, functions g(·) which have flat regions may lead to discrete/mixed RVs.

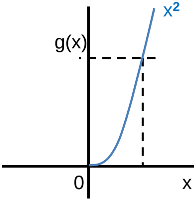
Derived Random Variables: Example

Let

$$g(x) = \begin{cases} x^2 & , x \geq 0 \\ 0 & , \text{otherwise} \end{cases}$$

and define $Y = g(X)$. Find $f_Y(y)$, $F_Y(y)$ given $f_X(x)$, $F_X(x)$.

. . . .



$$\text{Find } F_Y(y) = \begin{cases} 0 & , \text{y} < 0 \\ F_X(0) & , \text{y} = 0 \\ ??? & , \text{y} > 0 \end{cases}$$

$$F_Y(0) = P[Y \leq 0] = P[Y = 0] = P[X \leq 0] = F_X(0)$$

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

$$= P[X \leq \sqrt{y}] = F_X(\sqrt{y})$$

Derived Random Variables: Example (cont.)

Let

$$g(x) = \begin{cases} x^2 & , x \geq 0 \\ 0 & , \text{otherwise} \end{cases}$$

and define $Y = g(X)$. Find $f_Y(y)$, $F_Y(y)$ given $f_X(x)$, $F_X(x)$.

. . . .

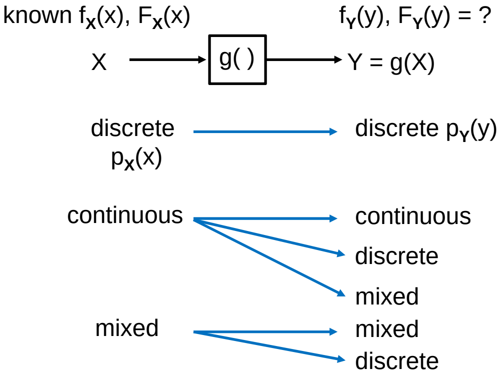
- Suppose $X \sim N(0,1)$. Then, $F_X(0) = 1/2$.

  Then, $F_Y(y) = \begin{cases} 0 & , y < 0 \\ 1/2 & , y = 0 \\ F_X(\sqrt{y}) & , y > 0 \end{cases}$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \begin{cases} \frac{1}{2}\delta(y) + \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) & , y \geq 0 \\ 0 & , \text{otherwise} \end{cases}$$

$$\text{where, } f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

Summary: Possibilities for Derived Distributions

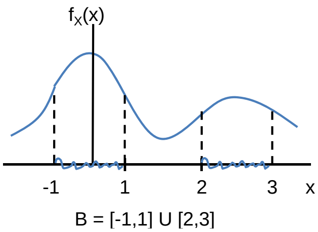


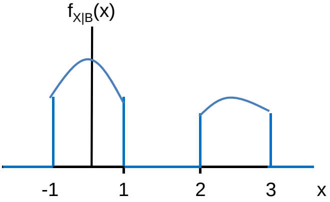
Conditioning a Continuous Random Variable

- Suppose that X has PDF $f_X(x)$ and let B be an event (i.e., a subset of $\mathbb{R}$, with $P[B] > 0$).
- **Definition:** The conditional PDF of X given B is given by

$$f_{X|B}(x) = \begin{cases} \frac{f_X(x)}{P[B]} & , x \in B \\ 0 & , \text{otherwise} \end{cases}$$

- **Interpretation:** Having observed B, we know that X must lie in this set, so the new PDF is the same as the old one, but renormalized by $P[B]$.





Conditioning a Continuous Random Variable: Conditional Expectations

$$E[X|B] = \int_{-\infty}^{+\infty} x f_{X|B}(x) dx$$

$$E[g(X)|B] = \int_{-\infty}^{+\infty} g(x) f_{X|B}(x) dx$$

Conditioning a Continuous Random Variable: Example

- Suppose that the holding time (duration) in minutes, T, of a telephone call is known to have an exponential distribution.
- $T \sim \exp(1/3)$ or

$$f_T(t) = \begin{cases} \frac{1}{3} e^{-1/3t} & , t \geq 0 \\ 0 & , \text{otherwise} \end{cases}$$

- Let $B = \{T > 2\}$. Find $f_{T|B}(t)$.

. . . .

$$\begin{aligned} P[B] = \int_2^{+\infty} f_T(t)dt &= 1 - P[T \leq 2] \\ &= 1 - (1 - e^{-2/3}) \\ &= e^{-2/3} \end{aligned}$$

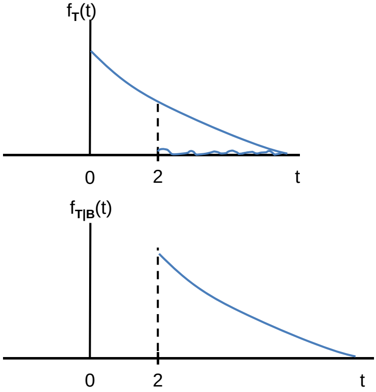
$$f_{T|B}(t) = \begin{cases} \frac{f_T(t)}{P[B]} = \frac{\frac{1}{3} e^{-1/3t}}{e^{-2/3}} = \frac{1}{3} e^{-\frac{1}{3}(t-2)} & , t > 2 \\ 0 & , \text{otherwise} \end{cases}$$

Conditioning a Continuous Random Variable: Example (cont.)

- Let $B = \{T > 2\}$. Find $E[T|B]$.

. . . .

$$
\begin{aligned}
E[T|B] &= \int_{-\infty}^{+\infty} t f_{T|B}(t)dt \\
&= \int_{2}^{+\infty} t \frac{1}{3} e^{-\frac{1}{3}(t-2)} dt \\
&= 5 \text{ minutes}
\end{aligned}
$$

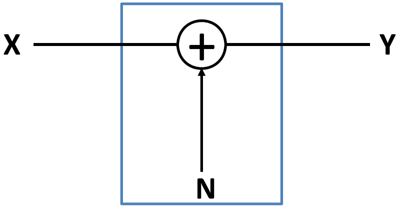


Reminder on integration by parts:

$$\int_a^b f(x)g'(x)dx = f(x)g(x)|_a^b - \int_a^b f'(x)g(x)dx$$

Multiple Continuous Random Variables

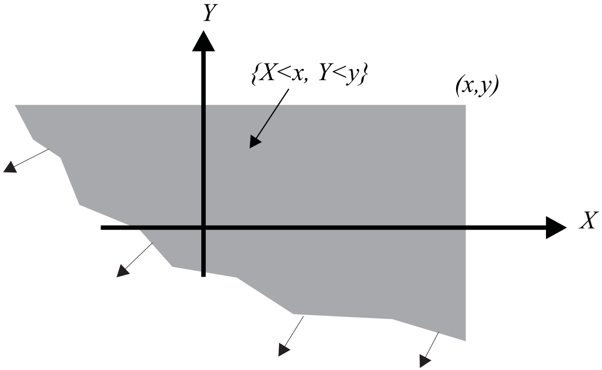
- **Example:** We would like to consider pairs of continuous RVs, e.g., (X, Y). Experiment produces at least two continuous RVs.



Joint CDF

- **Definition: (Joint CDF)** The joint CDF of X and Y is given by

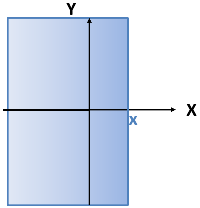
$$F_{X,Y}(x,y) = P[X \leq x, Y \leq y]$$



Multiple Continuous RVs: Joint CDF Properties

- $0 \leq F_{X,Y}(x,y) \leq 1$

- $$\begin{aligned} F_{X,Y}(x,+\infty) &= P[X \leq x, Y \leq +\infty] \\ &= P[X \leq x] \\ &= F_X(x) \end{aligned}$$



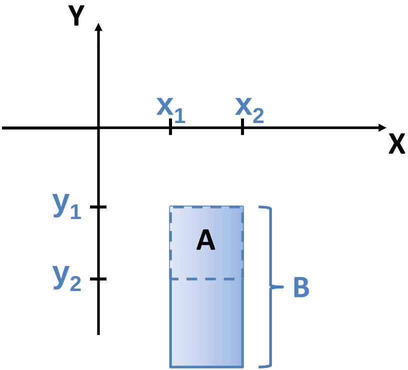
- $F_Y(y) = F_{X,Y}(+\infty, y)$

- $F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0$

- If $x_1 \geq x$ and $y_1 \geq y$, then
  $$F_{X,Y}(x_1, y_1) \geq F_{X,Y}(x,y).$$

Multiple Continuous RVs: Joint CDF and Rectangles

- We can use the joint CDF to compute the probability associated with rectangles as follows:



- $P[B] = F_{X,Y}(x_2, y_1) - F_{X,Y}(x_1, y_1)$
- $P[A] = P[B] - (F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2))$

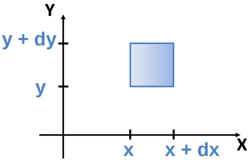
Joint Probability Density Function (PDF)

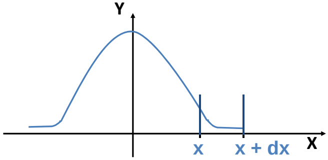
- **Definition: (Joint PDF)** The joint PDF of (X, Y) is $f_{X,Y}(x, y)$ satisfying

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v)dvdu$$

$$\text{equivalently, } f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$$

- **Interpretation:** $f_{X,Y}$ as the probability per unit area around (x, y). It can exceed 1, but must be such that $f_{X,Y} \geq 0$.
$P[x \leq X \leq x + dx, y \leq Y \leq y + dy] \approx f_{X,Y}(x, y)dxdy$





$$P[x \leq X \leq x + dx] \approx f_X(x)dx$$

Joint PDF Properties

- $f_{X,Y}(x, y) \geq 0$ (for all (x, y))
- $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dxdy = 1.$
- For any event $A \subset \mathbb{R}^2$ (i.e., subset of the x-y plane)

$$P[A] = \int_A \int f_{X,Y}(x, y) dxdy$$

Marginal PDF

- **Definition: (Marginal PDF)** Experiment produces continuous RVs X and Y, with joint PDF $f_{X,Y}(x,y)$, marginal PDFs are given by

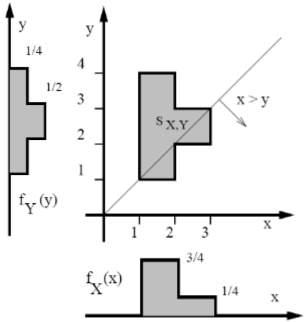
$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy$$
$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dx$$

- **Proof:** Write $F_X(x)$ as an integral, take the derivative.

Marginal PDF: Example

- Joint PDF which is uniform on region shown below.
- Find the constant c and marginals.

. . . .



Marginal PDF: Example

- Joint PDF which is uniform on region shown on previous slide.
- Find $P[X \geq Y]$.

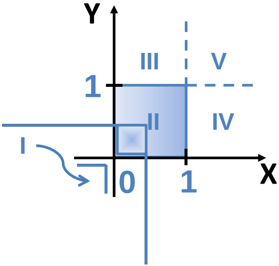
. . . .

- Let $B = \{(x, y) | x \geq y\}$

$$
\begin{aligned}
P[X \geq Y] = P[(X, Y) \in B] &= \int_B \int f_{X,Y}(x, y) dx dy \\
&= \frac{1}{4} Area(A \cap S_{X,Y}) = \frac{1}{4}
\end{aligned}
$$

Marginal PDF Example: Uniform Joint PDF

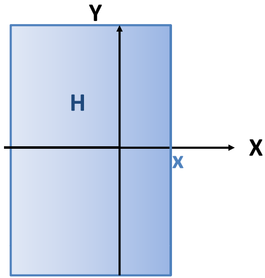
- Suppose (X, Y) is a randomly selected point out of the unit square.



Then, $f_{X,Y}(x,y) = \begin{cases} 1 & , 0 \le x, y \le 1 \\ 0 & , \text{otherwise} \end{cases}$

I: $F_{X,Y}(x,y) = 0$
II: $F_{X,Y}(x,y) = x \cdot y$ (x, y) are in region II: $0 \le x \le 1, 0 \le y \le 1$
III: $F_{X,Y}(x,y) = x$
IV: $F_{X,Y}(x,y) = y$
V: $F_{X,Y}(x,y) = 1$

Marginal CDF



$$
\begin{aligned}
F_X(x) &= P[X \leq x] \\
&= P[X \leq x, Y \leq \infty] \\
&= \int_H \int f_{X,Y}(\alpha, \beta) d\alpha d\beta \\
&= \int_{\alpha=-\infty}^{x} \int_{\beta=-\infty}^{\infty} f_{X,Y}(\alpha, \beta) d\beta d\alpha
\end{aligned}
$$

$$
f_X(x) = \int_{\beta=-\infty}^{\infty} f_{X,Y}(\alpha, \beta) d\beta
$$

Independent RVs

- X and Y are independent if $\forall x, y,\ F_{X,Y}(x,y) = F_X(x)F_Y(y)$ (equivalently, if $\forall x, y,\ f_{X,Y}(x,y) = f_X(x)f_Y(y)$).
- **Example:** Let X and Y be uniform on $[0,1] \times [0,1]$

$$f_{X,Y}(x,y) = \begin{cases} 1 & , 0 \le x \le 1, 0 \le y \le 1 \\ 0 & , \text{otherwise} \end{cases}$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \begin{cases} 1 & , 0 \le x \le 1 \\ 0 & , \text{otherwise} \end{cases}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \begin{cases} 1 & , 0 \le y \le 1 \\ 0 & , \text{otherwise} \end{cases}$$

Functions of Two Random Variables (I)

- **Example:** Receiver outputs X and Y from two antennas.

$$
\begin{aligned}
W_1 &= max(X, Y) \\
W_2 &= X + Y \\
W_3 &= aX + bY
\end{aligned}
$$

- What is the PDF of W_i?

....

- Find the CDF of W_i first.

$$
\begin{aligned}
F_{W_1}(w_1) &= P[W_1 \leq w_1] \\
&= P[max(X, Y) \leq w_1] \\
&= P[X \leq w_1, Y \leq w_1] \\
&= F_{X,Y}(w_1, w_1)
\end{aligned}
$$

Functions of Two Random Variables (II)

- If X and Y were independent, we could write $F_X(w_1)F_Y(w_1)$ instead of $F_{X,Y}(w_1, w_1)$:

$$F_{W_1}(w_1) = F_X(w_1)F_Y(w_1)$$
$$f_{W_1}(w_1) = f_X(w_1)F_Y(w_1) + F_X(w_1)f_Y(w_1)$$

[the derivative of the product of two functions]

- If X and Y are not independent

$$f_{W_1}(w_1) = \frac{\partial F_{X,Y}(w_1, w_1)}{\partial x}\Big|_{(w_1,w_1)} + \frac{\partial F_{X,Y}(w_1, w_1)}{\partial y}\Big|_{(w_1,w_1)}$$

Functions of Two Random Variables (III)

- If X and Y were independent

$$
\begin{aligned}
F_{W_2}(w_2) &= P[W_2 \leq w_2] \\
&= P[X + Y \leq w_2] \\
&= \int_A \int f_{X,Y}(x,y)dxdy \\
&= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{w_2-y} f_{X,Y}(x,y)dxdy \\
&= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{w_2-y} f_X(x)f_Y(y)dxdy \\
&= \int_{y=-\infty}^{\infty} f_Y(y)F_X(w_2-y)dy \\
f_{W_2}(w_2) &= \int_{y=-\infty}^{\infty} f_Y(y)f_X(w_2-y)dy \\
F_X(h(w_2)) &= f_X(h(w_2)) = f_X(w_2-y) \\
h(w_2) &= w_2 - y \\
f_{W_2} &= f_X * f_Y \text{ (Convolution)}
\end{aligned}
$$

Functions of Two Random Variables

- **Theorem:** For continuous random variables X and Y, the CDF of $W = g(X, Y)$ is

$$F_W(w) = P[W \leq w] = P[g(X, Y) \leq w] = \iint_{g(x,y) \leq w} f_{X,Y}(x, y)\,dxdy$$

$W = g(X, Y)$ Examples

- $W_1 = X + Y$
- $W_2 = max(X, Y)$
- $W_3 = XY$
- $W_4 = X/Y$

Finding the Expected Value $E[W]$

- We want to find the expectation of $W = g(X, Y)$. $(E[W] = E[g(X, Y)])$

- **Method 1:** Find the PDF of W, $f_W(w)$, then calculate

$$E[W] = \int_{-\infty}^{\infty} w f_W(w) dw$$

- **Method 2:** We can also compute the expected value of $W = g(X, Y)$ without going through the complicated process of deriving a probability model for W

$$E[W] = E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dxdy$$

Expectation of Sums

- Expected value of $g(X, Y) = g_1(X, Y) + ... + g_n(X, Y)$ is

$$E[g(X, Y)] = E[g_1(X, Y)] + ... + E[g_n(X, Y)]$$

- Sums:

$$\begin{aligned} E[X + Y] &= E[X] + E[Y] \\ E[aX + bY + c] &= aE[X] + bE[Y] + c \text{ (Linear operator)} \\ Var[X + Y] &= Var[X] + Var[Y] + 2Cov[X, Y] \end{aligned}$$

- Covariance:

$$Cov[X, Y] = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_X \mu_Y$$

Correlation Coefficient

- **Definition:** Correlation coefficient of two random variables X and Y is

$$\rho_{X,Y} = \frac{Cov[X,Y]}{\sqrt{Var[X]Var[Y]}}$$

- **Theorem:** $-1 \leq \rho_{X,Y} \leq 1$
- Same as for discrete random variables

Two Types of Conditioning

- By the occurrence of an event B of nonzero probability
  - Typically, this event B will be described in terms of a relationship between X and Y such as $X < Y$ or $X + Y \leq 100$.
  - Conditioning $f_{X,Y}(x, y)$ by an event is essentially the same as conditioning $f_X(x)$ by an event.
- By the observation that one of the random variables, say X, takes on the value x

Conditional Joint PDF

- When we learn that an event B occurs, we need to adjust our probability model for X and Y to reflect this knowledge.
- This modified probability model is the conditional joint PDF $f_{X,Y|B}(x, y)$.
- Given an event B with $P[B] > 0$, the conditional joint PDF of X and Y is

$$f_{X,Y|B}(x, y) = \begin{cases} \frac{f_{X,Y}(x,y)}{P[B]} & , (x, y) \in B \\ 0 & , \text{otherwise} \end{cases}$$

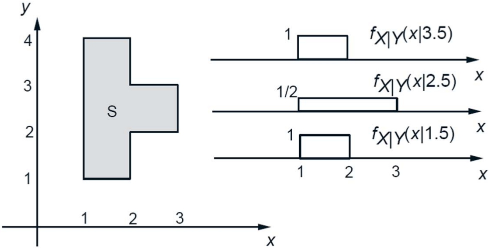
- Remove samples that do not belong to B and normalize.

Conditional PDF of Y Given $X = x$

- View joint PDF along slice $X = x$ and renormalize.
- $f_{Y|X}(y|x)$:
$$f_{Y|X}(y|x)dy = P[y \leq Y \leq y + dy | x \leq X \leq x + dx]$$
- Using Bayes' Theorem,
$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$
- $f_{X,Y}(x,y) = f_{Y|X}(y|x)f_X(x) = f_{X|Y}(x|y)f_Y(y)$

Conditional PDFs: Example

- For the joint PDF and $B = [0,2] \times [0,2]$, what do $f_{X,Y|B}$, $f_{X|Y}$, and $f_{Y|X}$ look like?



$$f_{X|Y}(x|3.5) = \frac{f_{X,Y}(x,3.5)}{f_Y(3.5)} = \frac{1/4}{1/4} = 1$$

$$f_{X|Y}(x|2.5) = \frac{f_{X,Y}(x,2.5)}{f_Y(2.5)} = \frac{1/4}{1/2} = 1/2$$

$$f_{X|Y}(x|1.5) = \frac{f_{X,Y}(x,1.5)}{f_Y(1.5)} = \frac{1/4}{1/4} = 1$$

Conditional Expected Value

- **Definition: (Conditional Expected Value)** If $f_Y(y) > 0$, the conditional expected value of X given $Y = y$ is

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

- **Definition: (Conditional Expected Value of a Function)** For any y such that $f_Y(y) > 0$, the conditional expected value of $g(X, Y)$ given $Y = y$ is

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

- Special case: conditional variance $Var[X|Y = y]$

$$Var[X|Y = y] = E[(X - E[X|Y = y])^2|Y = y]$$

Expected Value of Conditional Expected Value

- Note that the conditional expected value $E[g(X, Y)|Y = y]$ is a function of the observed value y of random variable Y.
- We can view the conditional expected value as a function of the random variable Y denoted $E[g(X, Y)|Y]$.
- Since $E[g(X, Y)|Y]$ is a function of Y, it is a random variable.
- We calculate the expected value of $E[g(X, Y)|Y]$ just as we would for any function $h(Y)$.
- **Theorem:**

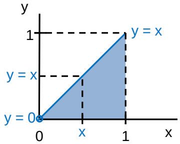
$$E[E[g(X, Y)|Y]] = \int_{-\infty}^{\infty} E[g(X, Y)|Y = y] f_Y(y) dy = E[g(X, Y)]$$

Expected Values: Example

- Let X and Y be random variables with joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 6y & , 0 \leq y \leq x \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

- Find the marginal PDF $f_X(x)$.



- For $0 \leq x \leq 1$,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_0^x 6ydy = 3x^2$$

- So,

$$f_X(x) = \begin{cases} 3x^2 & , 0 \leq x \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

Expected Values: Example (cont.)

- Let X and Y be random variables with joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 6y & , 0 \leq y \leq x \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

- Find the conditional PDF $f_{Y|X}(y|x)$. For what values of x is $f_{Y|X}(y|x)$ defined?

. . . .

- $f_{Y|X}(y|x)$ defined wherever $f_X(x) > 0$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \begin{cases} 2y/x^2 & , 0 \leq y \leq x \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

Expected Values: Example (cont.)

- Let X and Y be random variables with joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 6y & , 0 \leq y \leq x \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

- Find the conditional expected value $E[Y|X=x]$.

. . . .

$$E[Y|X=x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy \quad = \int_0^x y \frac{2y}{x^2} dy = \frac{2}{x^2} \left[\frac{y^3}{3}\right]_0^x = \frac{2}{3}x$$

Independent Continuous RVs

- **Definition: (Independence)** Continuous RVs X and Y are independent iff:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

- If X and Y are independent,

$$f_{X|Y}(x|y) = f_X(x) \quad f_{Y|X}(y|x) = f_Y(y)$$

Independence: Example 1

- Are X and Y independent?

$$f_{X,Y}(x,y) = \begin{cases} 4xy & , 0 \le x \le 1,\ 0 \le y \le 1 \\ 0 & , \text{otherwise} \end{cases}$$

. . . .

- The marginal PDFs of X and Y are

$$f_X(x) = \begin{cases} 2x & , 0 \le x \le 1 \\ 0 & , \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} 2y & , 0 \le y \le 1 \\ 0 & , \text{otherwise} \end{cases}$$

- Is $f_{X,Y}(x,y) = f_X(x) f_Y(y)$ for all pairs (x,y)? Yes. X and Y are independent.

Independence: Example 2

- Are U and V independent?

$$f_{U,V}(u,v) = \begin{cases} 24uv & , u \geq 0,\ v \geq 0,\ u+v \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

....

- Region of nonzero density is triangular and

$$f_U(u) = \begin{cases} 12u(1-u)^2 & , 0 \leq u \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

$$f_V(v) = \begin{cases} 12v(1-v)^2 & , 0 \leq v \leq 1 \\ 0 & , \text{otherwise} \end{cases}$$

- Is $f_{U,V}(u,v) = f_U(u)f_V(v)$? No. U and V are not independent!
- Learning the value of U changes our knowledge of V.
- For example, learning that $U = 1/2$ informs us that the event $P[V \leq 1/2] = 1$.

Independence: Example Summary

- In these two examples, we see that the region of nonzero probability plays a crucial role in determining whether random variables are independent.

Properties of Independent Continuous RVs

- **Theorem:** For independent random variables X and Y

$$
\begin{aligned}
E[g(X)h(Y)] &= E[g(X)]E[h(Y)] \\
Cov[X, Y] &= 0 \\
Var[X + Y] &= Var[X] + Var[Y]
\end{aligned}
$$

Jointly Gaussian Random Variables

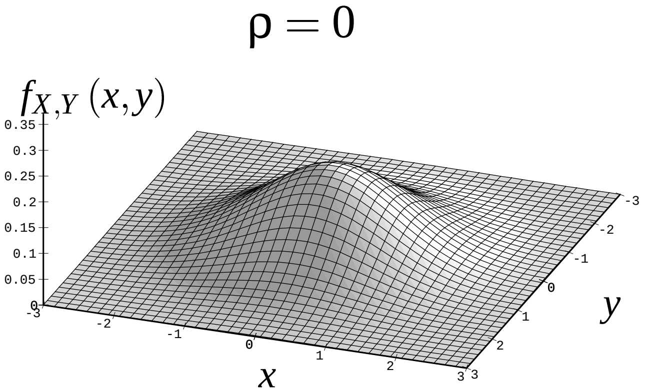
- **Definition:** X and Y have a bivariate Gaussian PDF if

$$f_{X,Y}(x,y) = \frac{exp\left[-\frac{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2}{2(1-\rho^2)}\right]}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

where $\sigma_1 > 0$, $\sigma_2 > 0$, $-1 \leq \rho \leq 1$

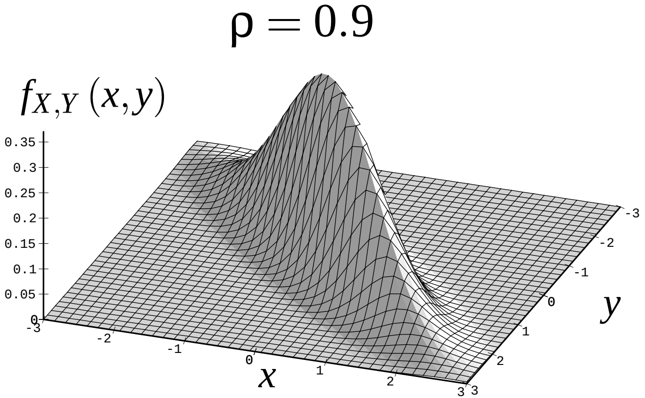
When $\rho = 0$

- $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$
- Joint PDF has circular symmetry of a hat



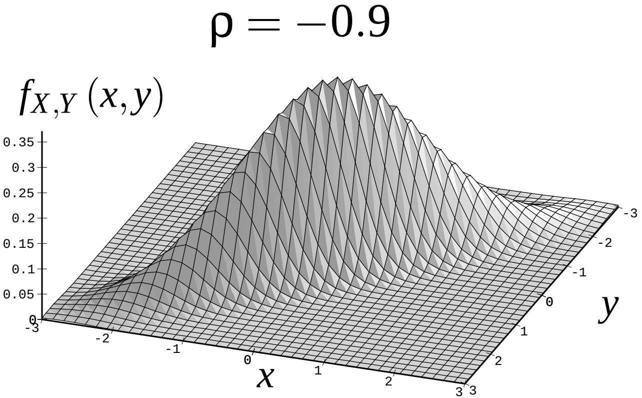
When $\rho = 0.9$

- $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$
- Joint PDF forms a ridge over the line $x = y$
- The ridge becomes increasingly steep as $\rho \to 1$



When $\rho = -0.9$

- $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$
- Joint PDF forms a ridge over the line $x = -y$
- The ridge becomes increasingly steep as $\rho \to -1$



Rewriting the Bivariate Gaussian PDF

- Complete the square of the exponent to write

$$f_{X,Y}(x,y) = f_X(x) f_{Y|X}(y|x)$$

where

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(x-\mu_1)^2/2\sigma_1^2}$$
$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\tilde{\sigma}_2} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2}$$

Bivariate Gaussian Properties

- $E[X] = \mu_1$
- Given $X = x$, Y is Gaussian
- Conditional mean of Y given $X = x$:

$$\begin{aligned}\tilde{\mu}_2(x) &= \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1)\\ &= E[Y|X = x]\end{aligned}$$

Gaussian Marginal PDF When $\rho = 0$ (X and Y are Uncorrelated)

- **Theorem:** If X and Y are the bivariate Gaussian random variables in our definition above and $\rho = 0$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(x-\mu_1)^2/2\sigma_1^2}$$
$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-(y-\mu_2)^2/2\sigma_2^2}$$

Gaussian Conditional PDF

- Given the marginal PDFs of X and Y, we use the definition of the conditional PDF to find the conditional PDFs.
- If X and Y are the bivariate Guassian random variables defined above, the conditional PDF of Y given X is

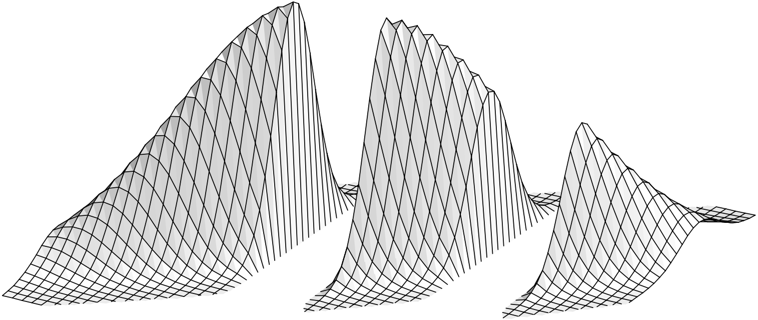
$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\tilde{\sigma}_2} e^{-(y-\tilde{\mu}_2(x))^2/2\tilde{\sigma}_2^2}$$

where

$$\tilde{\mu}_2(x) = E[Y|X = x] = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1)$$
$$\tilde{\sigma}_2^2 = Var[Y|X = x] = \sigma_2^2(1 - \rho^2)$$

Gaussian Conditional PDF

- Cross-sectional view of the joint Gaussian PDF with $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, and $\rho = 0.9$
- The bell shape of the cross section occurs because the conditional PDF $f_{Y|X}(y|x)$ is Gaussian.



More Than Two Continuous RVs

- **Definition: (Multivariate Joint CDF)** The joint CDF of $X_1, ..., X_n$ is

$$F_{X_1,...,X_n}(x_1, ..., x_n) = P[X_1 \leq x_1, ..., X_n \leq x_n]$$

- **Definition: (Multivariate Joint PDF)** The joint PDF of $X_1, ..., X_n$ is $f_{X_1,...,X_n}(x_1, ..., x_n)$ satisfying

$$F_{X_1,...,X_n}(x_1, ..., x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{X_1,...,X_n}(u_1, ..., u_n) du_1 ... du_n$$

Joint PDF Properties

- $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \frac{\partial^n F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)}{\partial x_1 \ldots \partial x_n}$
- $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) \geq 0$
- $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) dx_1 \ldots dx_n = 1$
- $P[A] = \underset{A}{\int \cdots \int} f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) dx_1 dx_2 \ldots dx_n$

Marginal PDFs

- **Theorem:** For a joint PDF of four random variables, $f_{W,X,Y,Z}(w,x,y,z)$, some marginal PDFs are

$$
\begin{aligned}
f_{X,Y,Z}(x,y,z) &= \int_{-\infty}^{\infty} f_{W,X,Y,Z}(w,x,y,z)dw \\
f_{W,Z}(w,z) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{W,X,Y,Z}(w,x,y,z)dxdy \\
f_{X}(x) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{W,X,Y,Z}(w,x,y,z)dwdydz
\end{aligned}
$$

- Can be generalized in a straightforward way to any marginal PDF of a joint PDF of an arbitrary number of random variables.

N Independent Random Variables

- **Definition: (N Independent Random Variables)** $X_1, ..., X_n$ are independent if

$$f_{X_1,...,X_n}(x_1, ..., x_n) = f_{X_1}(x_1)...f_{X_n}(x_n)$$

for all $x_1, ..., x_n$.

N Independent Random Variables

- Mutual independence of n random variables is typically the results of an experiment with special structure that ensures the independence
- The most common example occurs when an experiment consists of n independent trials.
- In this case, trial i produces the random variable X_i. Since all trials follow the same experiment, all of X_i have the same PDF. In this case, we say the random variables X_i are identically distributed.
- **Definition: (Independent and Identically Distributed)**
$X_1, ..., X_n$ are independent and identically distributed (iid) if and only if

$$f_{X_1,...,X_n}(x_1, ..., x_n) = f_X(x_1)...f_X(x_n)$$

for all $x_1, ..., x_n$.

Function of N Random Variables

- Just as we did for one and two random variables, we can derive a new random variable $Y = g(X_1, ..., X_n)$ that is a function of n random variables.

- When the X_i are continuous, we can find the CDF of Y

$$F_Y(y) = P[Y \leq y] = \int \cdots \int_{g(x_1,...,x_n) \leq y} f_{X_1,...,X_n}(x_1, ..., x_n) dx_1 ... dx_n$$

Expectation of a Function of N Random Variables

- **Theorem:** For $Y = g(X_1, ..., X_n)$, the expected value is

$$
\begin{aligned}
E[Y] &= E[g(X_1, ..., X_n)] \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, ..., x_n) f_{X_1,...,X_n}(x_1, ..., x_n) dx_1 ... dx_n
\end{aligned}
$$

- When $(X_1, ..., X_n)$ are independent, the expected value of $g(X_1) \times \cdots \times g(X_n)$ is the product of the expected values.

- **Theorem:** If $X_1, ..., X_n$ are independent random variables,

$$
E[g(X_1, ..., X_n)] = E[g(X_1)] \cdots E[g(X_n)]
$$

N Random Variables: Example 1

- Let $X_1, ..., X_n$ be iid RVs, with mean 0, variance 1 and covariance $Cov[X_i, X_j] = \rho$.
- Find the expected value and variance of the sum $Y = X_1, ..., X_n$.

. . . .

- The mean value of a sum of random variables is always the sum of their individual means.

$$E[Y] = \sum_{i=1}^{n} E[X_i] = 0$$

N Random Variables: Example 1 (cont.)

- The variance of any sum of random variables can be expressed in terms of the individual variances and covariances.

- Since $E[Y]$ is zero, $Var[Y] = E[Y^2]$. Thus,

$$Var[Y] = E\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] = E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right]$$

$$= \sum_{i=1}^{n} E[X_i^2] + \sum_{i=1}^{n}\sum_{j \neq i}^{n} E[X_i X_j]$$

- Since $E[X_i] = 0$, $E[X_i^2] = Var[X_i] = 1$ and for $i \neq j$

$$E[X_i X_j] = Cov[X_i, X_j] = \rho$$

- Thus,

$$Var[Y] = n + n(n-1)\rho$$

N Random Variables: Example 2

- Let $X_1, ..., X_n$ denote n iid random variables each with PDF $f_X(x)$.
- Find the CDF and PDF of $Y = min(X_1, ..., X_n)$.

. . . .

N Random Variables: Example 2 (cont.)

- We have

$$\begin{aligned} P[Y \geq y] &= P[min(X_1, ..., X_n) \geq y] \\ &= P[X_1 \geq y, ..., X_n \geq y) \geq y] \\ &= (P[X_1 \geq y])^n \\ &= [1 - F_X(y)]^n \end{aligned}$$

- Therefore, the CDF is

$$\begin{aligned} F_Y(y) = P[Y \leq y] &= 1 - P[Y \geq y] \\ &= 1 - (1 - F_X(y))^n \end{aligned}$$

- So, the PDF is

$$f_Y(y) = \frac{dF_Y(y)}{dy} = n(1 - F_X(y))^{n-1} f_X(y)$$