

LECTURE 9: Conditioning on an event; Multiple continuous r.v.'s

- Conditioning a r.v. on an event
 - Conditional PDF
 - Conditional expectation and the expected value rule
 - Exponential PDF: memorylessness
 - Total probability and expectation theorems
 - Mixed distributions
- Jointly continuous r.v.'s and joint PDFs
 - From the joints to the marginals
 - Uniform joint PDF example
 - The expected value rule and linearity of expectations
 - The joint CDF

Conditional PDF, given an event

$$p_X(x) = \mathbf{P}(X = x) \quad f_X(x) \cdot \delta \approx \mathbf{P}(x \leq X \leq x + \delta)$$

$$p_{X|A}(x) = \mathbf{P}(X = x \mid A) \quad f_{X|A}(x) \cdot \delta \approx \mathbf{P}(x \leq X \leq x + \delta \mid A)$$

$$\mathbf{P}(X \in B) = \sum_{x \in B} p_X(x) \quad \mathbf{P}(X \in B) = \int_B f_X(x) dx$$

$$\mathbf{P}(X \in B \mid A) = \sum_{x \in B} p_{X|A}(x) \quad \mathbf{P}(X \in B \mid A) = \int_B f_{X|A}(x) dx$$

$$\sum_x p_{X|A}(x) = 1 \quad \int f_{X|A}(x) dx = 1$$

Conditional PDF of X , given that $X \in A$

$$\mathbf{P}(x \leq X \leq x + \delta \mid X \in A) \approx f_{X|X \in A}(x) \cdot \delta$$

$$f_{X|X \in A}(x) = \begin{cases} 0, & \text{if } x \notin A \\ \frac{f_X(x)}{\mathbf{P}(A)}, & \text{if } x \in A \end{cases}$$

Conditional expectation of X , given an event

$$\mathbb{E}[X] = \sum_x x p_X(x)$$

$$\mathbb{E}[X] = \int x f_X(x) dx$$

$$\mathbb{E}[X | A] = \sum_x x p_{X|A}(x)$$

$$\mathbb{E}[X | A] = \int x f_{X|A}(x) dx$$

Expected value rule:

$$\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$$

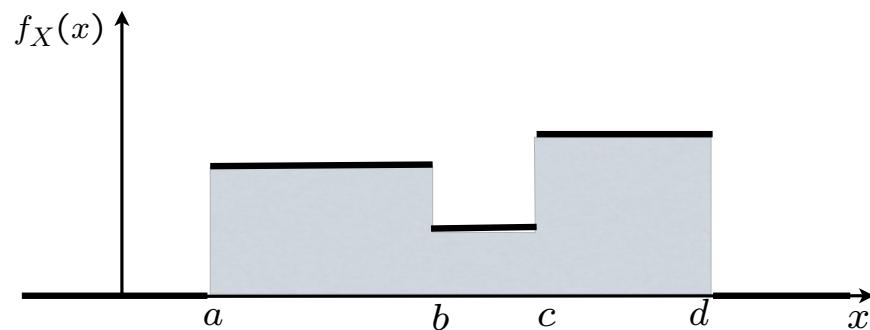
$$\mathbb{E}[g(X)] = \int g(x) f_X(x) dx$$

$$\mathbb{E}[g(X) | A] = \sum_x g(x) p_{X|A}(x)$$

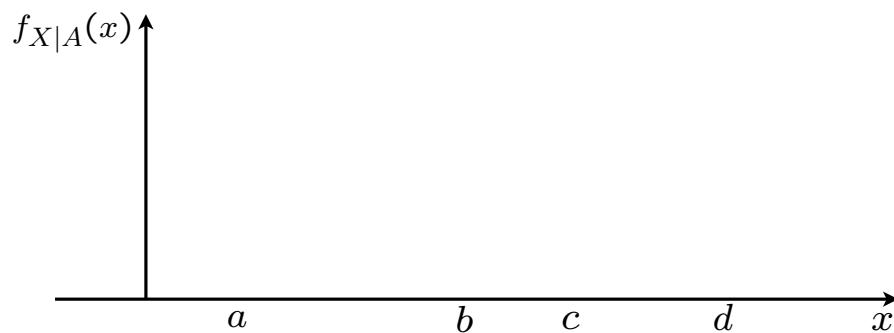
$$\mathbb{E}[g(X) | A] = \int g(x) f_{X|A}(x) dx$$

Example

$$A : \frac{a+b}{2} \leq X \leq b$$



$$\mathbf{E}[X \mid A] =$$



$$\mathbf{E}[X^2 \mid A] =$$

Memorylessness of the exponential PDF

- Do you prefer a used or a new “exponential” light bulb? **Probabilistically identical!**
- Bulb lifetime T : $\text{exponential}(\lambda)$

$$\mathbf{P}(T > x) = e^{-\lambda x}, \text{ for } x \geq 0$$

- we are told that $T > t$
- r.v. X : remaining lifetime

$$\mathbf{P}(X > x | T > t) = e^{-\lambda x}, \text{ for } x \geq 0$$

Memorylessness of the exponential PDF

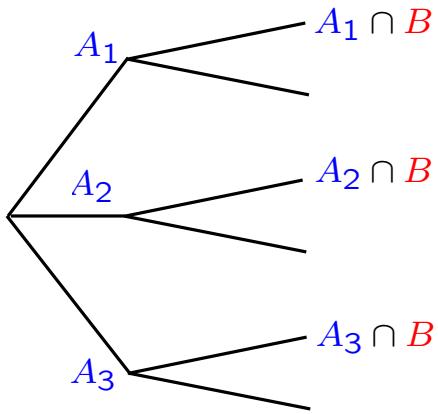
$$f_T(x) = \lambda e^{-\lambda x}, \quad \text{for } x \geq 0$$

$$\mathbf{P}(0 \leq T \leq \delta)$$

$$\mathbf{P}(t \leq T \leq t + \delta \mid T > t)$$

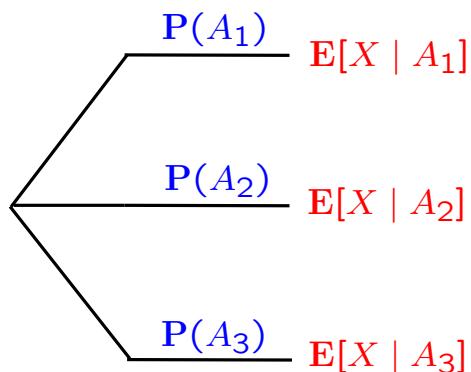
similar to an independent coin flip,
every δ time steps,
with $\mathbf{P}(\text{success}) \approx \lambda\delta$

Total probability and expectation theorems



$$P(B) = P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n)$$

$$p_X(x) = P(A_1)p_{X|A_1}(x) + \cdots + P(A_n)p_{X|A_n}(x)$$

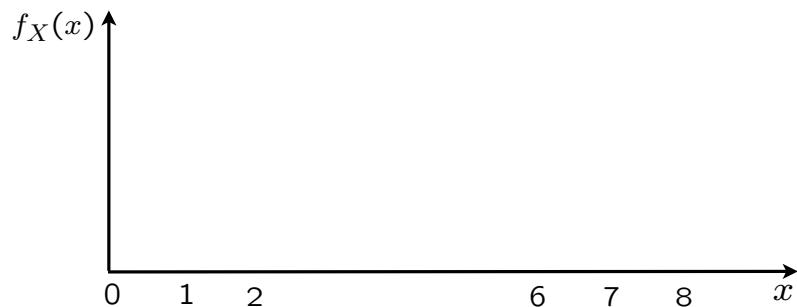


$$f_X(x) = P(A_1)f_{X|A_1}(x) + \cdots + P(A_n)f_{X|A_n}(x)$$

$$E[X] = P(A_1)E[X|A_1] + \cdots + P(A_n)E[X|A_n]$$

Example

- Bill goes to the supermarket shortly, with probability $1/3$, at a time uniformly distributed between 0 and 2 hours from now; or with probability $2/3$, later in the day at a time uniformly distributed between 6 and 8 hours from now



$$f_X(x) = \mathbf{P}(A_1)f_{X|A_1}(x) + \cdots + \mathbf{P}(A_n)f_{X|A_n}(x)$$

$$\mathbf{E}[X] = \mathbf{P}(A_1)\mathbf{E}[X | A_1] + \cdots + \mathbf{P}(A_n)\mathbf{E}[X | A_n]$$

Mixed distributions

$$X = \begin{cases} \text{uniform on } [0, 2], & \text{with probability } 1/2 \\ 1, & \text{with probability } 1/2 \end{cases}$$

Is X discrete?
Is X continuous?

Y discrete $X = \begin{cases} Y, & \text{with probability } p \\ Z, & \text{with probability } 1 - p \end{cases}$

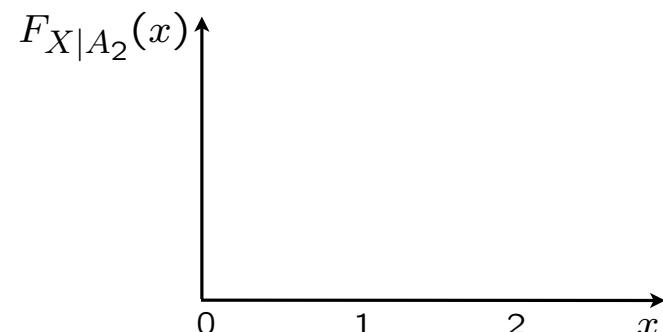
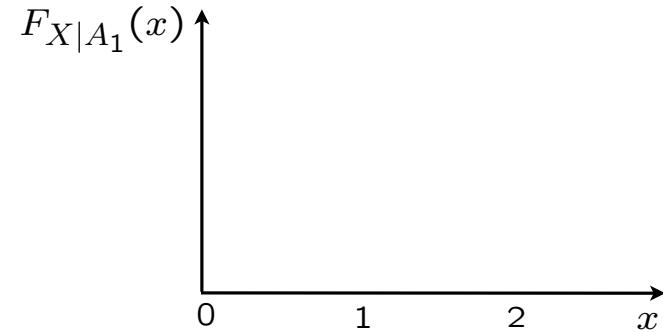
X is mixed

$$F_X(x) =$$

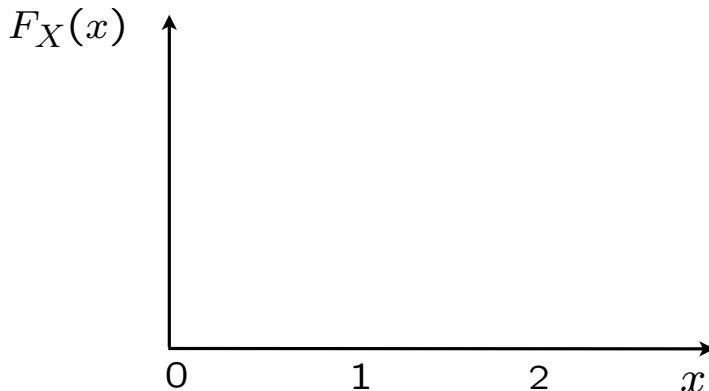
$$\mathbf{E}[X] =$$

Mixed distributions

$$X = \begin{cases} \text{uniform on } [0, 2], & \text{with probability } 1/2 \\ 1, & \text{with probability } 1/2 \end{cases}$$



$$F_X(x) = \mathbf{P}(A_1)F_{X|A_1}(x) + \mathbf{P}(A_2)F_{X|A_2}(x)$$



Jointly continuous r.v.'s and joint PDFs

$$\begin{array}{ll} p_X(x) & f_X(x) \\ p_{X,Y}(x,y) & f_{X,Y}(x,y) \end{array}$$

$$p_{X,Y}(x,y) = \mathbf{P}(X = x \text{ and } Y = y) \geq 0$$

$$f_{X,Y}(x,y) \geq 0$$

$$\mathbf{P}((X,Y) \in B) = \sum_{(x,y) \in B} p_{X,Y}(x,y)$$

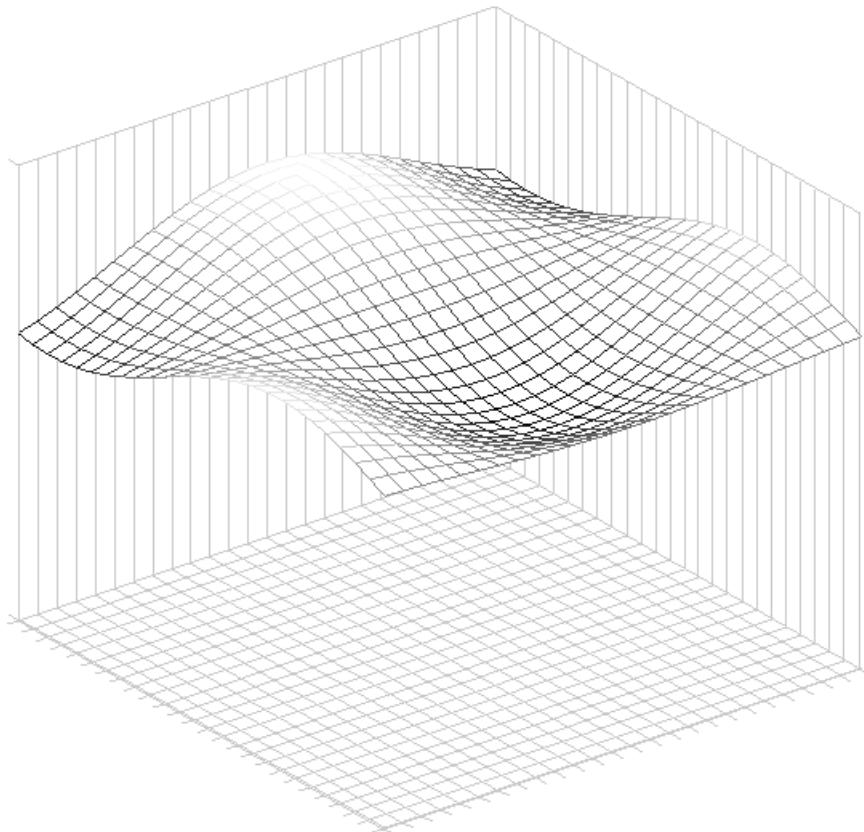
$$\mathbf{P}((X,Y) \in B) = \iint_{(x,y) \in B} f_{X,Y}(x,y) dx dy$$

$$\sum_x \sum_y p_{X,Y}(x,y) = 1$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx dy = 1$$

Definition: Two random variables are **jointly continuous** if they can be described by a joint PDF

Visualizing a joint PDF



$$\mathbf{P}((X, Y) \in B) = \iint_{(x,y) \in B} f_{X,Y}(x, y) dx dy$$

On joint PDFs

$$\mathbf{P}((X, Y) \in B) = \iint_{(x,y) \in B} f_{X,Y}(x, y) dx dy$$

$$\mathbf{P}(a \leq X \leq b, c \leq Y \leq d) = \int_c^d \int_a^b f_{X,Y}(x, y) dx dy$$

$$\mathbf{P}(a \leq X \leq a + \delta, c \leq Y \leq c + \delta) \approx f_{X,Y}(a, c) \cdot \delta^2$$

$f_{X,Y}(x, y)$: probability per unit area

$$\text{area}(B) = 0 \Rightarrow \mathbf{P}((X, Y) \in B) = 0$$

From the joint to the marginals

$$p_X(x) = \sum_y p_{X,Y}(x,y)$$

$$f_X(x) = \int f_{X,Y}(x,y) dy$$

$$p_Y(y) = \sum_x p_{X,Y}(x,y)$$

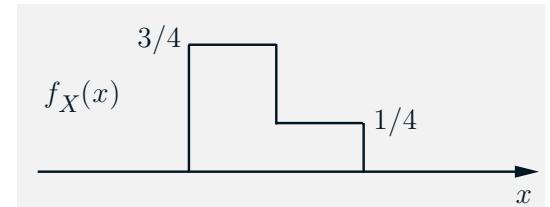
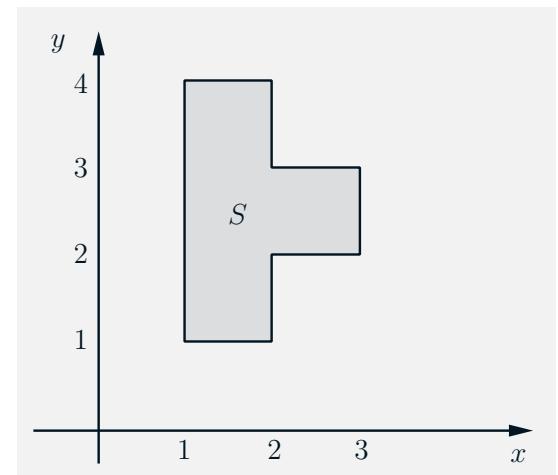
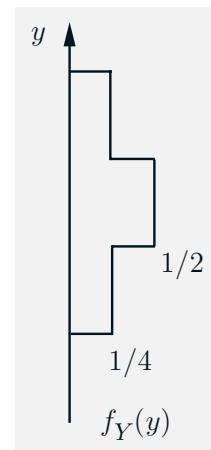
$$f_Y(y) = \int f_{X,Y}(x,y) dx$$

Uniform joint PDF on a set S

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\text{area of } S}, & \text{if } (x,y) \in S, \\ 0, & \text{otherwise.} \end{cases}$$

$$f_X(x) = \int f_{X,Y}(x,y) dy$$

$$f_Y(y) = \int f_{X,Y}(x,y) dx$$



More than two random variables

$$p_{X,Y,Z}(x, y, z)$$

$$f_{X,Y,Z}(x, y, z)$$

$$\sum_x \sum_y \sum_z p_{X,Y,Z}(x, y, z) = 1$$

$$p_X(x) = \sum_y \sum_z p_{X,Y,Z}(x, y, z)$$

$$p_{X,Y}(x, y) = \sum_z p_{X,Y,Z}(x, y, z)$$

Functions of multiple random variables

$$Z = g(X, Y)$$

Expected value rule:

$$\mathbb{E}[g(X, Y)] = \sum_x \sum_y g(x, y) p_{X,Y}(x, y)$$

$$\mathbb{E}[g(X, Y)] = \int \int g(x, y) f_{X,Y}(x, y) dx dy$$

Linearity of expectations

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$$\mathbb{E}[X_1 + \dots + X_n] = \mathbb{E}[X_1] + \dots + \mathbb{E}[X_n]$$

The joint CDF

$$F_X(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^x f_X(t) dt \quad f_X(x) = \frac{dF_X}{dx}(x)$$

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

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