THE UNIVERSITY OF TEXAS AT AUSTIN
Department of Statistics and Data Sciences
College of Natural Sciences

SDS 321: Introduction to Probability and Statistics

Lecture 16: Continuous random variables- Standardization and Joint distributions

Purnamrita Sarkar
Department of Statistics and Data Science
The University of Texas at Austin

www.cs.cmu.edu/∼psarkar/teaching

Roadmap

- Standardizing a Normal
- To read a normal table
- Joint PDF
- Marginal PDF
- Conditional PDF

The standard normal

- It is often helpful to map our normal distribution with mean μ and variance σ^2 onto a normal distribution with mean 0 and variance 1.
- This is known as the **standard normal**
- If we know probabilities associated with the standard normal, we can use these to calculate probabilities associated with normal random variables with arbitary mean and variance.
- If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{x - \mu}{\sigma} \sim N(0, 1)$.
- (Note, we often use the letter Z for standard normal random variables)

The standard normal

- I tell you that, if $X \sim N(0, 1)$, then $P(X < -1) = 0.159$.
- If $Y \sim N(1, 1)$, what is $P(Y < 0)$?
- Well we need to use the table of the **Standard Normal**.
- How do I transform Y such that it has the standard normal distribution?
- We know that a linear function of a normal random variable is also normally distributed!

The standard normal

- I tell you that, if $X \sim N(0,1)$, then $P(X < -1) = 0.159$.
- If $Y \sim N(1,1)$, what is $P(Y < 0)$?
- Well we need to use the table of the **Standard Normal**.
- How do I transform Y such that it has the standard normal distribution?
- We know that a linear function of a normal random variable is also normally distributed!
- Well $Z = Y - 1$ has mean zero and variance 1.

The standard normal

- I tell you that, if $X \sim N(0, 1)$, then $P(X < -1) = 0.159$.
- If $Y \sim N(1, 1)$, what is $P(Y < 0)$?
- Well we need to use the table of the **Standard Normal**.
- How do I transform Y such that it has the standard normal distribution?
- We know that a linear function of a normal random variable is also normally distributed!
- Well $Z = Y - 1$ has mean zero and variance 1.
- So $P(Y < 0) = P(Z < -1) = 0.159$.

The standard normal

- If $Y \sim N(0, 4)$, what value of y satisfies $P(Y < y) = 0.159$?
- The variance of Y is 4 times that of a standard normal random variable.
- Transform into a $N(0, 1)$ random variable!

The standard normal

- If $Y \sim N(0, 4)$, what value of y satisfies $P(Y < y) = 0.159$?
- The variance of Y is 4 times that of a standard normal random variable.
- Transform into a $N(0, 1)$ random variable!
- Use $Z = Y/2$...Now $Z \sim N(0, 1)$.

The standard normal

- If $Y \sim N(0, 4)$, what value of y satisfies $P(Y < y) = 0.159$?
- The variance of Y is 4 times that of a standard normal random variable.
- Transform into a $N(0, 1)$ random variable!
- Use $Z = Y/2$...Now $Z \sim N(0, 1)$.
- So, if $P(Y < y) = P(2Z < y) = P(Z < y/2)$.
- We want y such that $P(Z < y/2) = 0.159$. But we know that $P(Z < -1) = 0.159$, so?

The standard normal

- If $Y \sim N(0,4)$, what value of y satisfies $P(Y < y) = 0.159$?
- The variance of Y is 4 times that of a standard normal random variable.
- Transform into a $N(0,1)$ random variable!
- Use $Z = Y/2$...Now $Z \sim N(0,1)$.
- So, if $P(Y < y) = P(2Z < y) = P(Z < y/2)$.
- We want y such that $P(Z < y/2) = 0.159$. But we know that $P(Z < -1) = 0.159$, so?
- So $y/2 = -1$ and as a result $y = -2$...!

The standard normal

- The CDF of the standard normal is denoted Φ:

$$\Phi(z) = P(Z \leq z) = P(Z < z) = \frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{z} e^{-t^2/2} dt$$

- We cannot calculate this analytically.
- The **standard normal table** lets us look up values of $\Phi(y)$ for $y \geq 0$

	.00	.01	.02	0.03	0.04	$\cdots$
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	$\cdots$
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	$\cdots$
0.2	0.5793	**0.5832**	0.5871	0.5910	0.5948	$\cdots$
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	

$$P(Z < 0.21) = 0.5832$$

CDF of a normal random variable

If $X \sim N(3, 4)$, what is $P(X < 0)$?

- First we need to **standardize**:

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 3}{2}$$

- So, a value of $x = 0$ corresponds to a value of $z = -1.5$
- Now, we can translate our question into the standard normal:

$$P(X < 0) = P(Z < -1.5) = P(Z \leq -1.5)$$

- Problem... our table only gives $\Phi(z) = P(Z \leq z)$ for $z \geq 0$.

CDF of a normal random variable

If $X \sim N(3, 4)$, what is $P(X < 0)$?

- First we need to **standardize**:

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 3}{2}$$

- So, a value of $x = 0$ corresponds to a value of $z = -1.5$
- Now, we can translate our question into the standard normal:

$$P(X < 0) = P(Z < -1.5) = P(Z \leq -1.5)$$

- Problem... our table only gives $\Phi(z) = P(Z \leq z)$ for $z \geq 0$.
- But, $P(Z \leq -1.5) = P(Z \geq 1.5)$, due to symmetry.

CDF of a normal random variable

If $X \sim N(3, 4)$, what is $P(X < 0)$?

- First we need to **standardize**:

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 3}{2}$$

- So, a value of $x = 0$ corresponds to a value of $z = -1.5$
- Now, we can translate our question into the standard normal:

$$P(X < 0) = P(Z < -1.5) = P(Z \leq -1.5)$$

- Problem... our table only gives $\Phi(z) = P(Z \leq z)$ for $z \geq 0$.
- But, $P(Z \leq -1.5) = P(Z \geq 1.5)$, due to symmetry.
- Our table only gives us "less than" values.

CDF of a normal random variable

If $X \sim N(3, 4)$, what is $P(X < 0)$?

- First we need to **standardize**:

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 3}{2}$$

- So, a value of $x = 0$ corresponds to a value of $z = -1.5$
- Now, we can translate our question into the standard normal:

$$P(X < 0) = P(Z < -1.5) = P(Z \leq -1.5)$$

- Problem... our table only gives $\Phi(z) = P(Z \leq z)$ for $z \geq 0$.
- But, $P(Z \leq -1.5) = P(Z \geq 1.5)$, due to symmetry.
- Our table only gives us "less than" values.
- But, $P(Z \geq 1.5) = 1 - P(Z < 1.5) = 1 - P(Z \leq 1.5) = 1 - \Phi(1.5)$.

CDF of a normal random variable

If $X \sim N(3, 4)$, what is $P(X < 0)$?

- First we need to **standardize**:

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 3}{2}$$

- So, a value of $x = 0$ corresponds to a value of $z = -1.5$
- Now, we can translate our question into the standard normal:

$$P(X < 0) = P(Z < -1.5) = P(Z \leq -1.5)$$

- Problem... our table only gives $\Phi(z) = P(Z \leq z)$ for $z \geq 0$.
- But, $P(Z \leq -1.5) = P(Z \geq 1.5)$, due to symmetry.
- Our table only gives us "less than" values.
- But, $P(Z \geq 1.5) = 1 - P(Z < 1.5) = 1 - P(Z \leq 1.5) = 1 - \Phi(1.5)$.
- And we're done!
  $P(X < 0) = 1 - \Phi(1.5) =$ (look at the table...)$1 - 0.9332 = 0.0668$

Recap

- With continuous random variables, any specific value of $X = x$ has zero probability.
- So, writing a function for $P(X = x)$ – like we did with discrete random variables – is pretty pointless.
- Instead, we work with **PDFs** $f_X(x)$ – functions that we can integrate over to get the probabilities we need.

$$P(X \in B) = \int_B f_X(x)dx$$

- We can think of the PDF $f_X(x)$ as the "probability mass per unit area" near x.
- We are often interested in the probability of $X \leq x$ for some x – we call this the cumulative distribution function $F_X(x) = P(X \leq x)$.
- Once we know $f_X(x)$, we can calculate expectations and variances of X.

Multiple continuous random variables

- Let X and Y be two continuous random variables.
- Each one takes on values on the real line, i.e. $X \in \mathbb{R}$ and $Y \in \mathbb{R}$.
- Together, each possible pair of values describe a point in the real plane, i.e. $(X, Y) \in \mathbb{R}^2$.
- We say X and Y are **jointly continous** if the probability of them jointly taking on values in some subset B of the plane can be described as

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

  using some continuous function $f_{X,Y}$, for all $B \in \mathbb{R}^2$ – i.e. all subsets of the 2-D plane.
- Notation means "integrate over all values of x and y s.t. $(x, y) \in B$

Joint PDF

- We call $f_{X,Y}$ the **joint pdf** of X and Y.
- It allows us to calculate the probability of any set of combinations of X and Y
  - e.g. the probability that a person weighs over 200lb and is under 6'
  - e.g. the probability that a person's height in inches is more than twice their weight in pounds.

  - So, this could describe the first scenario above, $P(200 \leq X \leq \infty, -\infty \leq Y \leq 6)$
  - In this case B is a rectangle
- What is $\int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} f_{X,Y}(x,y) dx\, dy$?

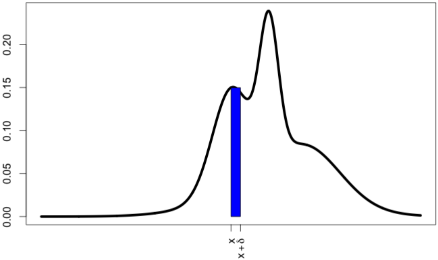
Joint PDF

- We call $f_{X,Y}$ the **joint pdf** of X and Y.
- It allows us to calculate the probability of any set of combinations of X and Y
  - e.g. the probability that a person weighs over 200lb and is under 6'
  - e.g. the probability that a person's height in inches is more than twice their weight in pounds.

  - So, this could describe the first scenario above, $P(200 \leq X \leq \infty, -\infty \leq Y \leq 6)$
  - In this case B is a rectangle
- What is $\int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} f_{X,Y}(x,y)dx\,dy$? 1

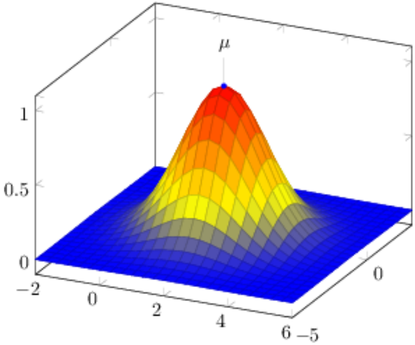
Joint PDF: Intuition

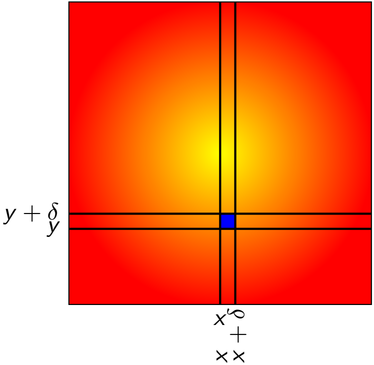
- Remember we could think of $f_X(x)$ as the "probability mass per unit length" near to x?



- Because $f_X(x) = \dfrac{P(x \leq X \leq x + \delta)}{\delta}$

Joint PDF: Intuition





- We can think of the joint PDF $f_{X,Y}(x, y)$ as the "probability mass per unit area" for a small area near X.
- Again, remember, $f_{X,Y}(x, y)$ **is not a probability!**

Multiple random variables to a single random variable

- We can get from the **joint PMF** of X and Y to the **marginal PMF** of X by summing over (marginalizing over) Y:

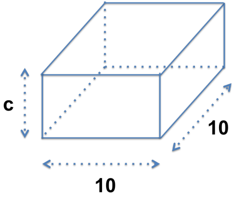
$$p_X(x) = \sum_y p_{X,Y}(x,y)$$

- We can get from the **joint PDF** of X and Y to the **marginal PDF** of X by integrating over (marginalizing over) Y:

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

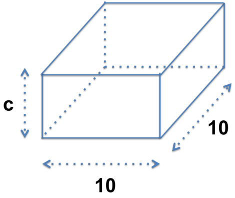
Example: Bivariate uniform random variable

- Anita (X) and Benjamin (Y) both pick a number between 0 and 10, according to a continuous uniform distribution. What is $f_{X,Y}(x, y)$?



Example: Bivariate uniform random variable

- Anita (X) and Benjamin (Y) both pick a number between 0 and 10, according to a continuous uniform distribution. What is $f_{X,Y}(x,y)$?



- Let's see... we know all pairs (x,y) are equally likely, so we know $f_{X,Y} = c$. It must satisfy $\int_{x=0}^{10}\int_{y=0}^{10} f_{X,Y}(x,y)dx\,dy = 1$.
- So, $c\underbrace{\int_{x=0}^{10}\int_{y=0}^{10} dx\,dy}_{100} = 1\ldots$
- So $c = f_{X,Y}(x,y) = 0.01$ for all $0 \leq x, y \leq 10$.

Example: marginal probability

- $f_{X,Y}(x,y) = \begin{cases} 0.01 & \text{If } x, y \in [0,10] \\ 0 & \text{otherwise} \end{cases}$
- What is $f_X(x)$?

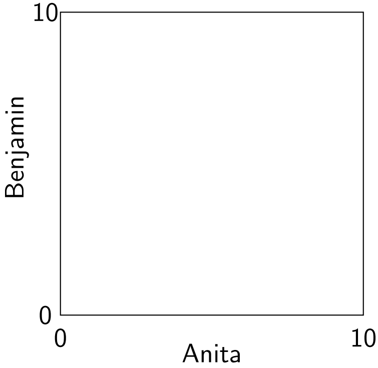
- In general, we will have $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$
- We have **marginalized out** one of our random variables... just like we did when looking at PMFs.
- We call $f_X(x)$ the **marginal PDF** of X

Example: marginal probability

- $f_{X,Y}(x,y) = \begin{cases} 0.01 & \text{If } x, y \in [0, 10] \\ 0 & \text{otherwise} \end{cases}$
- What is $f_X(x)$?
- $f_X(x) = \begin{cases} \int_{y=0}^{10} 0.01 dy = 0.1 & \text{If } x \in [0, 10] \\ 0 & \text{otherwise} \end{cases}$
- Not surprisingly $X \sim \mathit{Uniform}([0, 10])$ and $Y \sim \mathit{Uniform}([0, 10])$.
- In general, we will have $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$
- We have **marginalized out** one of our random variables... just like we did when looking at PMFs.
- We call $f_X(x)$ the **marginal PDF** of X

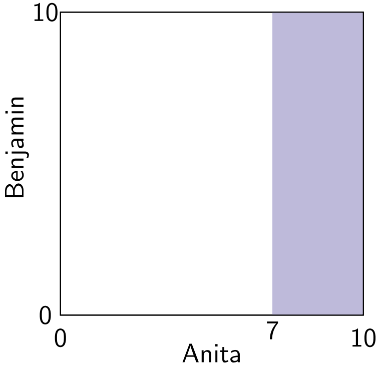
Example: marginalization

- What is the probability that Anita picks a number greater than 7?



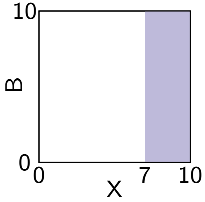
Example: marginalization

- What is the probability that Anita picks a number greater than 7?



- That's going to correspond to the shaded region... $P(X > 7) = 0.01(3 \times 10) = 0.3$.
- Or, using calculus: $\int_{x=7}^{10} \int_{y=0}^{10} f_{X,Y}(x, y) dx\, dy$

Marginalization

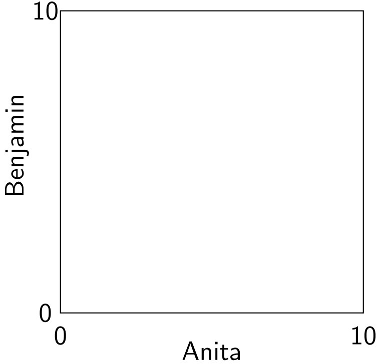


- $P(X > 7) = \int_{x=7}^{10} \int_{y=0}^{10} f_{X,Y}(x, y) dx\, dy$

- But, this doesn't depend on Benjamin at all! It is the same as $P(X > 7) = \int_{x>7} f_X(x) dx$.

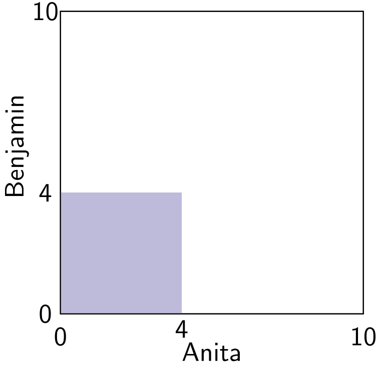
Example: Uniform random variable

- What is the probability that they both pick numbers less than 4?



Example: Uniform random variable

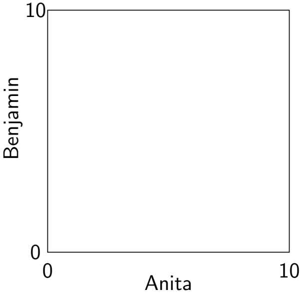
- What is the probability that they both pick numbers less than 4?



- It will be $0.01 \int_0^4 \int_0^4 dx\, dy = 0.01 \times 16 = 0.16$
  - i.e. $0.01 \times$ the shaded area.
  - Or $16/100$!

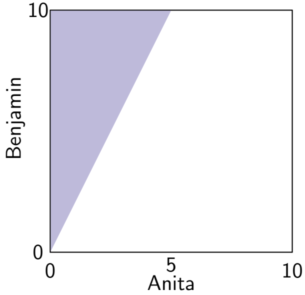
Example: Uniform random variable

- What is the probability that Benjamin picks a number at least twice that of Anita?



Example: Uniform random variable

- What is the probability that Benjamin picks a number at least twice that of Anita?

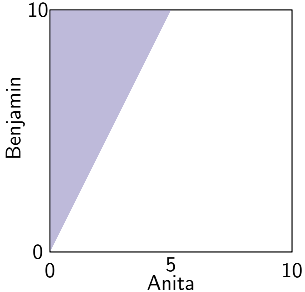


- That's going to correspond to the shaded region...
  $P(Y \geq 2X) = 0.01(0.5 \times 5 \times 10) = 0.25.$

- Or, using calculus: $\int_{x=0}^{10} \int_{y=2x}^{10} f_{X,Y}(x,y) dx\, dy = \int_{x=0}^{10} \int_{y=2x}^{10} c \times 1_{0 \leq x \leq 10, 0 \leq y \leq 10} dx\, dy$

Example: Uniform random variable

- What is the probability that Benjamin picks a number at least twice that of Anita?



- That's going to correspond to the shaded region...
  $P(Y \geq 2X) = 0.01(0.5 \times 5 \times 10) = 0.25$.
- Or, using calculus: $\int_{x=0}^{10} \int_{y=2x}^{10} f_{X,Y}(x,y) dx\, dy = \int_{x=0}^{10} \int_{y=2x}^{10} c \times 1_{0 \leq x \leq 10, 0 \leq y \leq 10} dx\, dy$
- $$\int_{x=0}^{10} \int_{y=2x}^{10} c \times 1_{0 \leq 2x \leq 10} dx\, dy = c \int_{0}^{5} dx = c \int_{x=0}^{5} (10 - 2x) dx$$
  $$= c(10 \times 5 - (5^2 - 0)) = 0.01 \times 25 = 0.25$$

Recap

- Last time, we introduced the idea of continuous random variables and PDFs.
- A PDF is a function we can integrate over to get $P(X \in B) = \int_B f_X(x)dx$.
- We extended this to look at **joint PDFs** and **conditional PDFs**.
- We can borrow results from conditional probability and probabilities of intersections!
- But we need to be careful to remember, a PDF is **not** a probability...
- Next time, we will continue looking at continuous probability distributions.