

# Unit 5 Continuous Random Variables

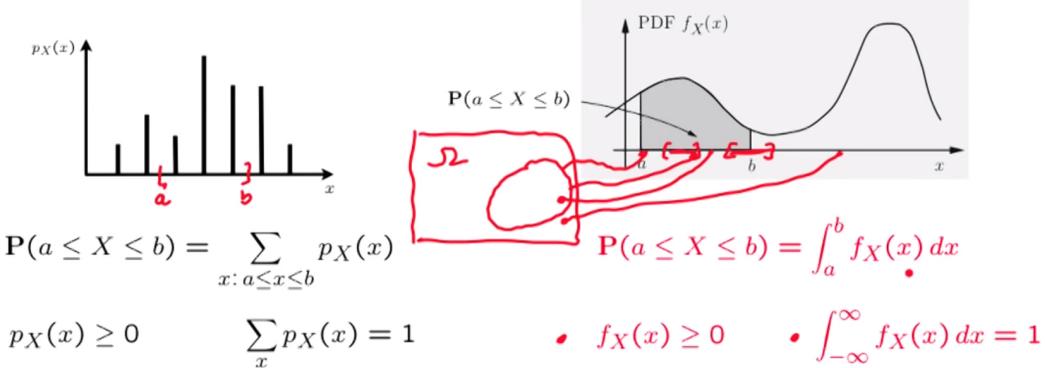
Very much a repetition of unit 4 but with continuous variables instead of discrete variables

Will use more calculus rather than counting/sums

## Probability Density Functions (PDFs)

The mass of probability is spread out over a line, calculating the density of this gives us a probability of the interval

### Probability density functions (PDFs)



**Definition:** A random variable is **continuous if it can be described by a PDF**

$$\mathbb{P}(1 \leq X \leq 3 \text{ or } 4 \leq X \leq 5) = \mathbb{P}(1 \leq X \leq 3) + \mathbb{P}(4 \leq X \leq 5)$$

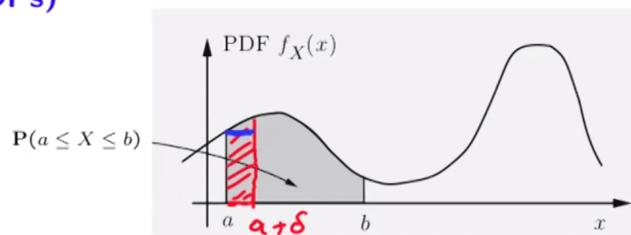
Any particular point has 0 probability but a group of points have a positive probability

## Probability density functions (PDFs)

$\delta > 0, \text{ small}$

$$\mathbb{P}(a \leq X \leq a + \delta)$$

$$\approx f_X(a) \cdot \delta$$



$$\mathbb{P}(a \leq X \leq a + \delta) \approx f_X(a) \cdot \delta$$

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x) dx$$

$$f_X(x) \geq 0$$

$$\int_{-\infty}^{\infty} f_X(x) dx = 1$$

$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(X=a) + \mathbb{P}(X=b) + \mathbb{P}(a < X < b)$$

## Exercise: PDFs

4/4 points (graded)

Let  $X$  be a continuous random variable with a PDF of the form

$$f_X(x) = \begin{cases} c(1-x), & \text{if } x \in [0, 1], \\ 0, & \text{otherwise.} \end{cases}$$

Find the following values.

1.  $c =$   ✓

2.  $\mathbb{P}(X = 1/2) =$   ✓

3.  $\mathbb{P}(X \in \{1/k : k \text{ integer}, k \geq 2\}) =$   ✓

4.  $\mathbb{P}(X \leq 1/2) =$   ✓

1. We have  $1 = \int_{-\infty}^{\infty} f_X(x) dx = \int_0^1 c(1-x) dx = c(x - x^2/2) \Big|_0^1 = c/2$ , and therefore,  $c = 2$ .

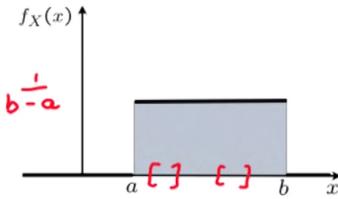
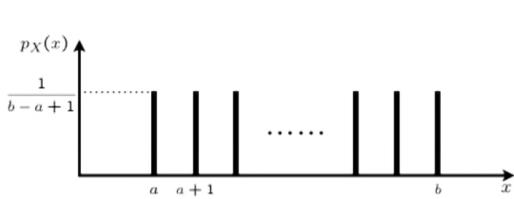
2. Individual points have zero probability.

3. Using countable additivity and the fact that single points have zero probability, we have

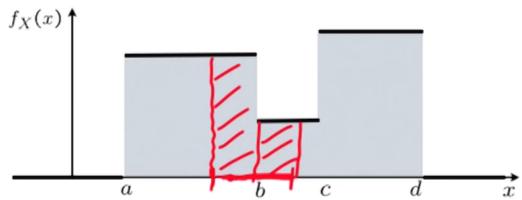
$$\mathbf{P}(X \in \{1/2, 1/3, 1/4, 1/5, \dots\}) = \sum_{n=2}^{\infty} \mathbf{P}(X = 1/n) = \sum_{n=2}^{\infty} 0 = 0.$$

4.  $\mathbf{P}(X \leq 1/2) = \int_{-\infty}^{1/2} f_X(x) dx = \int_0^{1/2} 2(1-x) dx = 2(x - x^2/2) \Big|_0^{1/2} = \frac{3}{4}$ .

### Example: continuous uniform PDF



- Generalization: piecewise constant PDF



### Exercise: Piecewise constant PDF

2/2 points (graded)

Consider a piecewise constant PDF of the form

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

Find the following values.

a)  $c =$   ✓ Answer: 0.25

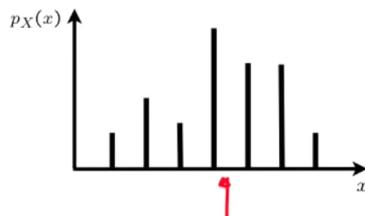
b)  $\mathbf{P}(1/2 \leq X \leq 3/2) =$   ✓ Answer: 0.375

#### Solution:

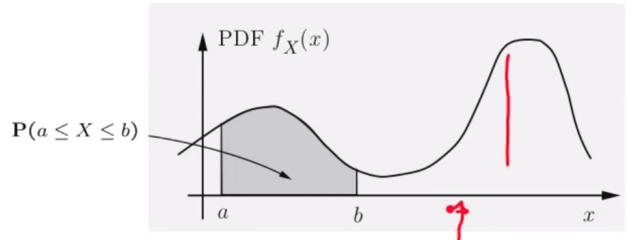
a) The total area under the PDF is the sum of the areas of two rectangles and is equal to  $(2c) \cdot 1 + c \cdot 2 = 4c$ . Therefore,  $c = 1/4$ .

b) The total area under the PDF over the interval of interest is the sum of the areas of two smaller rectangles and is equal to  $(2c) \cdot (1/2) + c \cdot (1/2) = c \cdot (3/2) = 3/8$ .

## Expectation/mean of a continuous random variable



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



$$\mathbf{E}[X] = \int_{-\infty}^{\infty} x \underline{f_X(x)} dx$$

- **Interpretation:** Average in large number of independent repetitions of the experiment

Fine print:  
Assume  $\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$

"Centre of gravity" = expectation

## Properties of expectations

- If  $X \geq 0$ , then  $\mathbf{E}[X] \geq 0$
- If  $a \leq X \leq b$ , then  $a \leq \mathbf{E}[X] \leq b$
- Expected value rule:

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

$$\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x) dx$$

$$\mathbf{E}[x^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx$$

- Linearity

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

## Variance and its properties

- Definition of variance:  $\text{var}(X) = E[(X - \mu)^2]$

$$\mu = E[X]$$

- Calculation using the expected value rule,  $E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x) dx$

$$\text{var}(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x) dx$$

$$g(x) = (x - \mu)^2$$

**Standard deviation:**  $\sigma_X = \sqrt{\text{var}(X)}$

✓  $\text{var}(aX + b) = a^2\text{var}(X)$

✓ A useful formula:  $\text{var}(X) = E[X^2] - (E[X])^2$

### Exercise: Uniform PDF

3/3 points (graded)

Let  $X$  be uniform on the interval  $[1, 3]$ . Suppose that  $1 < a < b < 3$ . Then,

(a)  $P(a \leq X \leq b) =$  (b-a)\*0.5 ✓ Answer:  $(b-a)/2$

(b - a) · 0.5

(Your answer to part (a) should be an algebraic expression involving  $a$  and  $b$ .)

(b)  $E[X] =$  2 ✓ Answer: 2

(c)  $E[X^3] =$  10 ✓ Answer: 10

#### Solution:

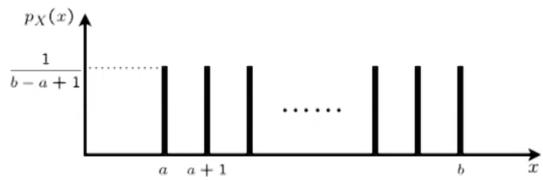
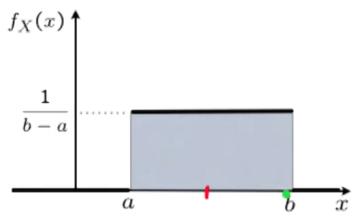
(a) The value of the PDF on the interval  $[1, 3]$  must be equal to  $1/2$ , so that it integrates to 1. Thus,

$$P(a \leq X \leq b) = \int_a^b \frac{1}{2} dx = \frac{b-a}{2}.$$

(b) The expected value of a uniform is the midpoint of its range:  $E[X] = (1 + 3)/2 = 2$ .

(c) Using the expected value rule,  $E[X^3] = \int_1^3 x^3 \cdot \frac{1}{2} dx = \frac{1}{2} \cdot \frac{1}{4} x^4 \Big|_1^3 = \frac{1}{2} \cdot \frac{1}{4} \cdot (81 - 1) = 10$ .

### Continuous uniform random variable; parameters $a, b$



$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx \\ = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{a+b}{2}$$

$$\mathbb{E}[X^2] = \int_a^b x^2 \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \left( \frac{b^3}{3} - \frac{a^3}{3} \right) \quad \text{var}(X) = \frac{1}{12}(b-a)(b-a+2)$$

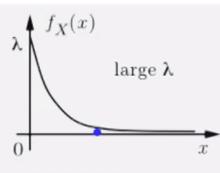
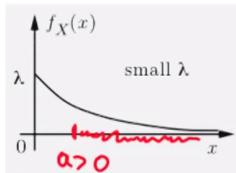
$$\text{var}(X) = E[X^2] - (E[X])^2 = \boxed{(b-a)^2/12} \quad \sigma = \frac{b-a}{\sqrt{12}}$$

### Exponential random variable; parameter $\lambda > 0$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$$\int f_X(x) dx = 1$$

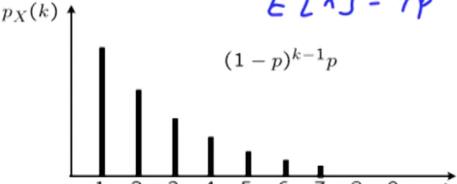
$$E[X] = 1/\lambda$$



$$\mathbb{E}[X] = \int_0^{\infty} x \cdot \lambda e^{-\lambda x} dx = 1/\lambda$$

$$\mathbb{E}[X^2] = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx = 2/\lambda^2$$

$$\text{var}(X) = E[X^2] - (E[X])^2 = 1/\lambda^2$$



$$\boxed{P(X \geq a)} = \int_a^{\infty} \lambda e^{-\lambda x} dx \\ \left[ \int e^{ax} dx = \frac{1}{a} e^{ax} \quad a \leftrightarrow -\lambda \right] \\ = \lambda \cdot \left( -\frac{1}{\lambda} \right) e^{-\lambda x} \Big|_a^{\infty} \\ = -e^{-\lambda \cdot \infty} + e^{-\lambda a} = \boxed{e^{-\lambda a}}$$

Shows that the probability decreases exponentially beyond  $a$ , and if  $a=0$  then the probability equals 1 which is the entire space of probability

When  $x$  is zero we have  $f(x) = \lambda$  since  $e^0 = 1$

The exponential is similar to the discrete geometric and are used in conjunction with one another

The exponential is also used a lot to model the time for an event to happen such as the time for a customer to arrive, a machine to break down, a meteorite to hit a house etc.

### Exercise: Exponential PDF

1/2 points (graded)

Let  $X$  be an exponential random variable with parameter  $\lambda = 2$ . Find the values of the following. Use 'e' for the base of the natural logarithm (e.g., enter  $e^{-3}$  for  $e^{-3}$ ).

a)  $E[(3X + 1)^2] =$  3.5 ✖ Answer: 8.5

b)  $P(1 \leq X \leq 2) =$  0.1170 ✓ Answer: 0.11702

**Solution:**

a) By expanding the quadratic, using linearity of expectations, and the facts that  $E[X] = 1/\lambda$  and  $E[X^2] = 2/\lambda^2$ , we have

$$E[(3X + 1)^2] = 9E[X^2] + 6E[X] + 1 = 9 \cdot \frac{2}{2^2} + 6 \cdot \frac{1}{2} + 1 = \frac{17}{2}.$$

b) We have seen that for  $a > 0$ , we have  $P(X \geq a) = e^{-\lambda a}$ , so that  $P(X \leq a) = 1 - e^{-\lambda a}$ . Therefore,

$$P(1 \leq X \leq 2) = P(X \leq 2) - P(X \leq 1) = (1 - e^{-4}) - (1 - e^{-2}) = e^{-2} - e^{-4}.$$

Will always use capital F to denote a CDF here

Once we know the CDF of a random variable we have enough information to calculate anything we need

Once we have the CDF we can get the PDF (via derivative of CDF) and vice versa (via integrating)

### Cumulative distribution function (CDF)

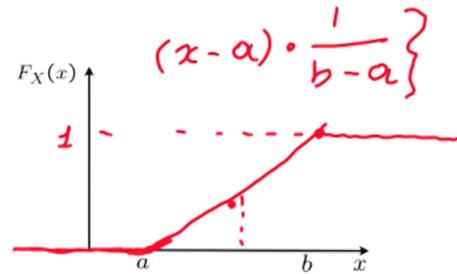
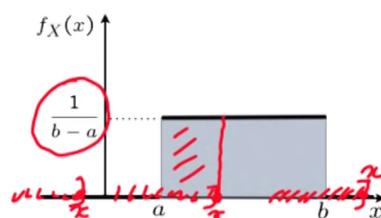
**CDF definition:**  $F_X(x) = P(X \leq x)$

- Continuous random variables:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t) dt$$

$$\underbrace{P(X \leq 4)}_{=} = \underbrace{P(X \leq 3)}_{=} + \underbrace{P(3 < X \leq 4)}_{=}$$

$$\boxed{\frac{dF_X(x)}{dx} = f_X(x)}$$

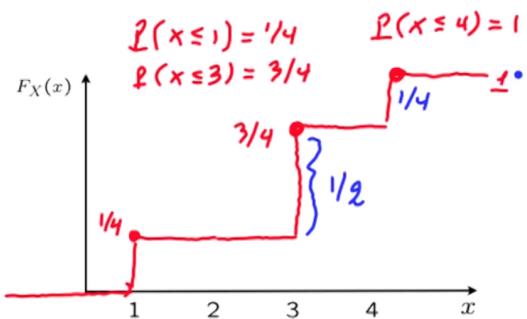
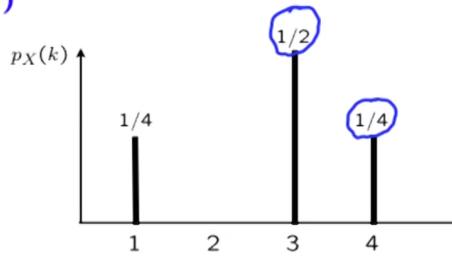


## Cumulative distribution function (CDF)

**CDF definition:**  $F_X(x) = P(X \leq x)$

- Discrete random variables:

$$F_X(x) = P(X \leq x) = \sum_{k \leq x} p_X(k)$$



## General CDF properties

$$F_X(x) = P(X \leq x)$$



- Non-decreasing      If  $y \geq x \Rightarrow F_X(y) \geq F_X(x)$
- $F_X(x)$  tends to 1, as  $x \rightarrow \infty$
- $F_X(x)$  tends to 0, as  $x \rightarrow -\infty$

### Exercise: Exponential CDF

1/2 points (graded)

Let  $X$  be an exponential random variable with parameter 2.

Find the CDF of  $X$ . Express your answer in terms of  $x$  using standard notation. Use 'e' for the base of the natural logarithm (e.g., enter  $e^{-3*x}$  for  $e^{-3x}$ ).

a) For  $x \leq 0$ ,  $F_X(x) =$   ✓ Answer: 0

0

b) For  $x > 0$ ,  $F_X(x) =$   ✗ Answer: 1-e<sup>-2\*x</sup>

1

**STANDARD NOTATION**

**Solution:**

a) Since  $X$  is a nonnegative random variable,  $F_X(x) = \mathbf{P}(X \leq x) = 0$  for  $x \leq 0$ .

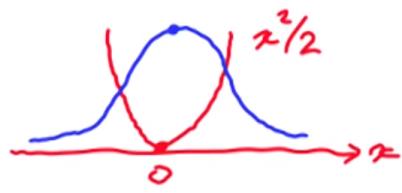
b) We have seen that for an exponential random variable with parameter  $\lambda$  and for any  $a > 0$ , we have  $\mathbf{P}(X \geq a) = e^{-\lambda a}$ . Therefore,  $F_X(x) = \mathbf{P}(X \leq x) = 1 - \mathbf{P}(X \geq x) = 1 - e^{-\lambda x} = 1 - e^{-2x}$ .

## Normal (Gaussian) random variables

- Important in the theory of probability
  - Central limit theorem
- Prevalent in applications
  - Convenient analytical properties
  - Model of noise consisting of many, small independent noise terms

## Standard normal (Gaussian) random variables

- Standard normal  $N(0, 1)$ :  $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$



calculus:

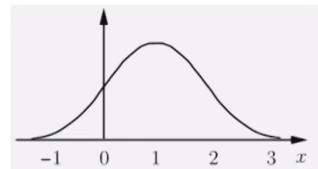
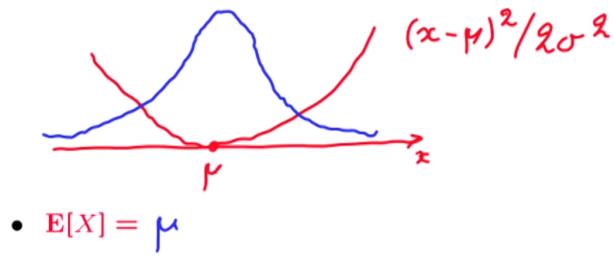
$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$$

- $E[X] = 0$
- $\text{var}(X) = 1$  integrate by parts

the  $1/\sqrt{2\pi}$  factor comes from calculus to ensure that the integral equals to 1

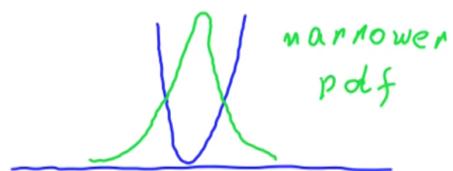
## General normal (Gaussian) random variables

- General normal  $N(\mu, \sigma^2)$ :  $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$   
 $\sigma > 0$



- $E[X] = \mu$
- $\text{var}(X) = \sigma^2$

$\sigma$  small



## Linear functions of a normal random variable

- Let  $Y = aX + b$   $X \sim N(\mu, \sigma^2)$

$$E[Y] = a\mu + b$$

$$\text{Var}(Y) = a^2 \sigma^2$$

- Fact (will prove later in this course):

$$Y \sim N(a\mu + b, a^2\sigma^2)$$

- Special case:  $a = 0$ ?

$$Y = b \quad \text{discrete}$$
$$N(b, 0)$$

N means Y is still a Normal random variable

### Exercise: Normal random variables

1/1 point (graded)

Choose the correct answer below.

According to our conventions, a normal random variable  $X \sim N(\mu, \sigma^2)$  is a continuous random variable

always.

if and only if  $\sigma \neq 0$ .

if and only if  $\mu \neq 0$  and  $\sigma \neq 0$ .



### Solution:

When  $\sigma \neq 0$ , the distribution of  $X$  is described by a PDF, and so  $X$  is a continuous random variable. But when  $\sigma = 0$ , then  $X$  has all of its probability assigned to a single point, and therefore it is not a continuous random variable. (For continuous random variables, any single point must have zero probability.)

## Calculation of normal probabilities

### Standard normal tables

- No closed form available for CDF

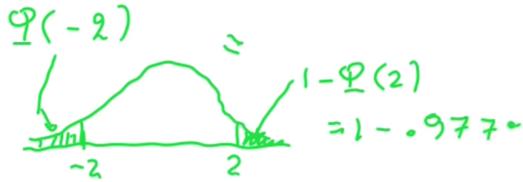
but have tables, for the standard normal

$$Y \sim N(0, 1)$$

$$\Phi(y) = F_Y(y) = P(Y \leq y)$$

$$\Phi(0) = P(Y \leq 0) = 0.5$$

$$\Phi(1.016) = 0.8770 \quad \Phi(2.9) = 0.9981$$



	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8794	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9227	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	Videos pos
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986

### Standardizing a random variable

- Let  $X$  have mean  $\mu$  and variance  $\sigma^2 > 0$

- Let  $Y = \frac{X - \mu}{\sigma}$

$$E[Y] = 0 \quad \text{Var}(Y) = \frac{1}{\sigma^2} \text{Var}(x) = 1$$

$$X = \mu + \sigma Y$$

- If also  $X$  is normal, then:  $Y \sim N(0, 1)$

$Y$  measures how many standard deviations it is from the mean

## Calculating normal probabilities

- Express an event of interest in terms of standard normal

$$X \sim N(6, 4) \quad \sigma = 2$$

*st. normal*

$$\frac{2 - 6}{2} \leq \frac{X - 6}{2} \leq \frac{8 - 6}{2}$$

$$\mathbb{P}(2 \leq X \leq 8) = \mathbb{P}(-2 \leq Y \leq 1)$$

$$= \mathbb{P}(Y \leq 1) - \mathbb{P}(Y \leq -2)$$

$$= \mathbb{P}(Y \leq 1) - (1 - \mathbb{P}(Y \leq 2))$$

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986

Standard normal because it is of the form  $X - \mu / \sigma$   
This gives  $Y$  so then we can use the standard normal tables

Exercise: Using the normal tables

3/3 points (graded)

Let  $X$  be a normal random variable with mean 4 and variance 9.

Use the [normal table](#) to find the following probabilities, to an accuracy of 4 decimal places.

Normal Table	
Show	
a) $\mathbb{P}(X \leq 5.2) =$	0.6554
	✓ Answer: 0.6554
b) $\mathbb{P}(X \geq 2.8) =$	0.6554
	✓ Answer: 0.6554
c) $\mathbb{P}(X \leq 2.2) =$	0.2743
	✓ Answer: 0.2743

**Solution:**

a) Note that the standard deviation is 3. Subtracting the mean and dividing by the standard deviation, we obtain

$$\mathbb{P}(X \leq 5.2) = \mathbb{P}\left(\frac{X - 4}{3} \leq \frac{5.2 - 4}{3}\right) = \Phi(0.4) = 0.6554.$$

b) Because of the symmetry around the mean,  $\mathbb{P}(X \geq 2.8) = \mathbb{P}(X \leq 5.2) = 0.6554$ .

$$\mathbb{P}(X \leq 2.2) = \mathbb{P}\left(\frac{X - 4}{3} \leq \frac{2.2 - 4}{3}\right) = \Phi(-0.6) = 1 - \Phi(0.6) = 1 - 0.7257 = 0.2743.$$

## Conditioning on an event; Multiple Continuous Random Variables

From discrete to continuous we replace sums with integrals and PMFs with PDFs

$$P(A) > 0$$

Conditional PDF, given an event

$$p_X(x) = P(X = x)$$

$$f_X(x) \cdot \delta \approx P(x \leq X \leq x + \delta)$$

$$p_{X|A}(x) = P(X = x | A)$$

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

$$P(X \in B) = \sum_{x \in B} p_X(x)$$

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

$$P(X \in B | A) = \sum_{x \in B} p_{X|A}(x)$$

$$P(X \in B | A) = \int_B f_{X|A}(x) dx$$

Def

$$\sum_x p_{X|A}(x) = 1$$

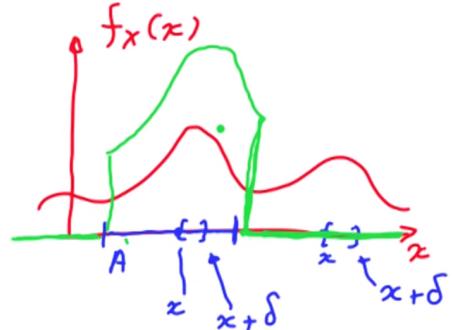
$$\int f_{X|A}(x) dx = 1$$

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

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

$$= \frac{P(x \leq X \leq x + \delta, X \in A)}{P(A)}$$

$$= \frac{P(x \leq X \leq x + \delta)}{P(A)} \approx \frac{f_X(x) \delta}{P(A)}$$



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

Rescaled so that the total probability under the conditional curve is equal to 1

## Conditional expectation of $X$ , given an event

$$\mathbf{E}[X] = \sum_x xp_X(x)$$

$$\mathbf{E}[X] = \int xf_X(x) dx$$

$$\mathbf{E}[X | A] = \sum_x xp_{X|A}(x)$$

$$\mathbf{E}[X | A] = \int xf_{X|A}(x) dx \quad \text{Def}$$

### Expected value rule:

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

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

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

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

### Exercise: A conditional PDF

0/1 point (graded)

Suppose that  $X$  has a PDF of the form

$$f_X(x) = \begin{cases} 1/x^2, & \text{if } x \geq 1, \\ 0, & \text{otherwise.} \end{cases}$$

For any  $x > 2$ , the conditional PDF of  $X$ , given the event  $X > 2$  is

( $1/(x^2)$ )/( $-1/x$ )

✖ Answer:  $2/(x^2)$

$$\frac{1}{x^2}$$
  
$$-\frac{1}{x}$$

(Your answer should be an algebraic function of  $x$ , in [standard notation](#).)

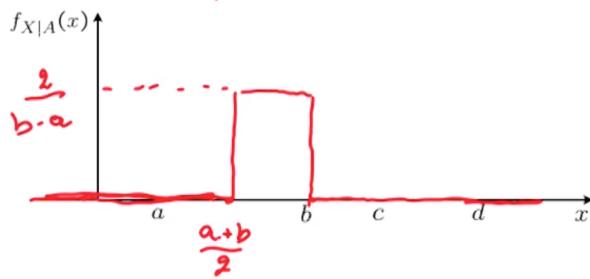
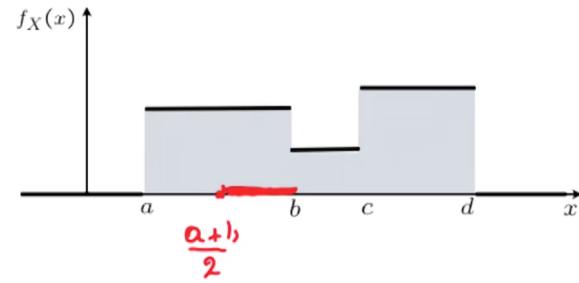
[STANDARD NOTATION](#)

#### Solution:

The conditional PDF will be a scaled version of the unconditional, of the form  $\frac{f_X(x)}{\mathbf{P}(X>2)}$ . Now,  $\mathbf{P}(X > 2) = \int_2^\infty \frac{1}{x^2} dx = -\frac{1}{x} \Big|_2^\infty = 1/2$ , and so the answer is  $2/x^2$ .

### Example

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



$$\begin{aligned} E[X | A] &= \frac{1}{2} \cdot \frac{a+b}{2} + \frac{1}{2} b \\ &= \frac{1}{4} a + \frac{3}{4} b \\ E[X^2 | A] &= \left. \frac{2}{b-a} \cdot x^2 dx \right|_{\frac{a+b}{2}}^{\frac{a+b}{2}} \end{aligned}$$

### Memorylessness of the exponential PDF

- Do you prefer a used or a new "exponential" light bulb? **Probabilistically identical!**
- Bulb lifetime  $T$ : exponential( $\lambda$ )
  - $P(T > x) = e^{-\lambda x}$ , for  $x \geq 0$ 
    - we are told that  $T > t$
    - r.v.  $X$ : remaining lifetime  $= T - t$



$$\begin{aligned} P(X > x | T > t) &= e^{-\lambda x}, \text{ for } x \geq 0 \\ &= \frac{P(T-t > x, T > t)}{P(T > t)} = \frac{P(T > t+x, T > t)}{P(T > t)} = \frac{P(T > t+x)}{P(T > t)} \\ &= \frac{e^{-\lambda(t+x)}}{e^{-\lambda t}} = e^{-\lambda x} \end{aligned}$$

This is saying that the probability of a lightbulb lasting  $x$  amount of time is the same for a new and a used lightbulb, the memorylessness of the probability ensures this

Electronic items, light bulbs etc., often exhibit the “no ageing” phenomenon. These items do not change properties with usage, but they fail when some external shock like a surge of high voltage, comes along. It can be shown that if these shocks occur according to a Poisson process then the lifetime of the item has exponential distribution.

Practical implications of no ageing :

(a) Since a used component is as good as new (stochastically), there is no advantage in following a policy of planned replacement of used components known to be still functioning.

### Memorylessness of the exponential PDF

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

$$\begin{aligned} P(0 \leq T \leq \delta) &\approx f_T(0) \cdot \delta = \lambda \delta \\ P(t \leq T \leq t + \delta | T > t) &= \approx \lambda \delta \end{aligned}$$



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

The probability of the lightbulb burning out in the next delta timestep is always the same

This is the foundation of the Poisson Process

## 6. Exercise: Memorylessness of the exponential

[Bookmark this page](#)

### Exercise: Memorylessness of the exponential

2/3 points (graded)

Let  $X$  be an exponential random variable with parameter  $\lambda$ .

a) The probability that  $X > 5$  is

b) The probability that  $X > 5$  given that  $X > 2$  is

c) Given that  $X > 2$ , and for a small  $\delta > 0$ , the probability that  $4 \leq X \leq 4 + 2\delta$  is approximately

**Solution:**

a) We have seen in the past that for an exponential random variable with parameter  $\lambda$ ,  $\mathbf{P}(X > a) = e^{-\lambda a}$ , and so  $\mathbf{P}(X > 5) = e^{-5\lambda}$ .

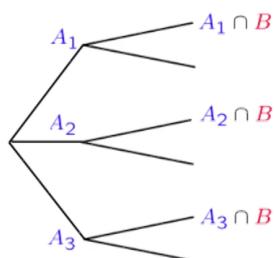
b) Because of the memorylessness property, given that  $X > 2$ , the remaining time  $X - 2$  is again exponential with the same parameter. Thus,  $\mathbf{P}(X > 5 | X > 2) = \mathbf{P}(X - 2 > 3 | X > 2) = \mathbf{P}(X > 3) = e^{-3\lambda}$ .

c) By memorylessness, this is the same as the unconditional probability that an exponential takes values in the interval  $[2, 2 + 2\delta]$ , which is approximately the length,  $2\delta$ , of the small interval times the density evaluated at 2, yielding  $2\lambda\delta e^{-2\lambda}$ .

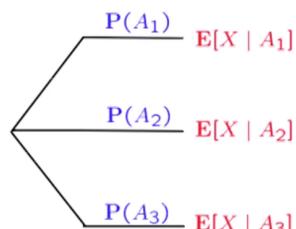
From lecture, we learned that  $\mathbf{P}(t < T < (t + \delta) | T > t) = \lambda\delta$

In c above the time change is  $2\delta$  not  $\delta$  so it is multiplied by  $2.\lambda.\delta$

### Total probability and expectation theorems



$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(A_1)\mathbf{P}(B | A_1) + \cdots + \mathbf{P}(A_n)\mathbf{P}(B | A_n) \\ p_X(x) &= \mathbf{P}(A_1)p_{X|A_1}(x) + \cdots + \mathbf{P}(A_n)p_{X|A_n}(x) \\ F_x(x) &= \mathbb{P}(X \leq x) = \mathbb{P}(A_1) \mathbb{P}(X \leq x | A_1) + \cdots \\ &\quad = \mathbb{P}(A_1) F_{x|A_1}(x) + \cdots\end{aligned}$$



$$\begin{aligned}f_X(x) &= \mathbf{P}(A_1)f_{X|A_1}(x) + \cdots + \mathbf{P}(A_n)f_{X|A_n}(x) \\ \int x f_x(x) dx &= \mathbb{P}(A_1) \int x f_{x|A_1}(x) dx + \cdots \\ \mathbf{E}[X] &= \mathbf{P}(A_1)\mathbf{E}[X | A_1] + \cdots + \mathbf{P}(A_n)\mathbf{E}[X | A_n]\end{aligned}$$

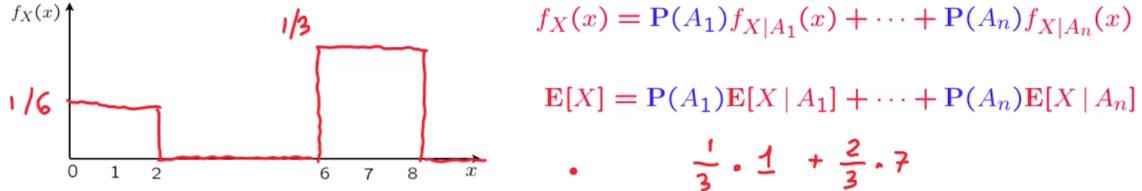
Derivative of CDF = PDF

Same formula as discrete case (lower) but now X is a continuous random variable

### 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

$$\mathbb{P}(A_1) = \frac{1}{3} \quad f_{X|A_1} \sim \text{unif}[0, 2] \quad \mathbb{P}(A_2) = \frac{2}{3} \quad f_{X|A_2} \sim U[6, 8]$$



### Exercise: Total probability theorem II

2/2 points (graded)

On any given day, mail gets delivered by either Alice or Bob. If Alice delivers it, which happens with probability  $1/4$ , she does so at a time that is uniformly distributed between 9 and 11. If Bob delivers it, which happens with probability  $3/4$ , he does so at a time that is uniformly distributed between 10 and 12. The PDF of the time  $X$  that mail gets delivered satisfies

a)  $f_X(9.5) =$  1/8 ✓ Answer: 0.125

b)  $f_X(10.5) =$  4/8 ✓ Answer: 0.5

#### Solution:

The PDF is  $1/4$  times a uniform on  $[9, 11]$  (of height  $1/2$ ) plus  $3/4$  times a uniform on  $[10, 12]$  (again of height  $1/2$ ).

a) At time 9.5, only the first uniform is nonzero, yielding  $f_X(9.5) = (1/4) \cdot (1/2) = 1/8$ .

b) At time 10.5 both uniforms are nonzero, yielding  $f_X(10.5) = (1/4) \cdot (1/2) + (3/4) \cdot (1/2) = 1/2$ .

## 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? **No**

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

Is  $X$  continuous? **No**  
 $P(X=1) = 1/2$   
 $X$  is mixed

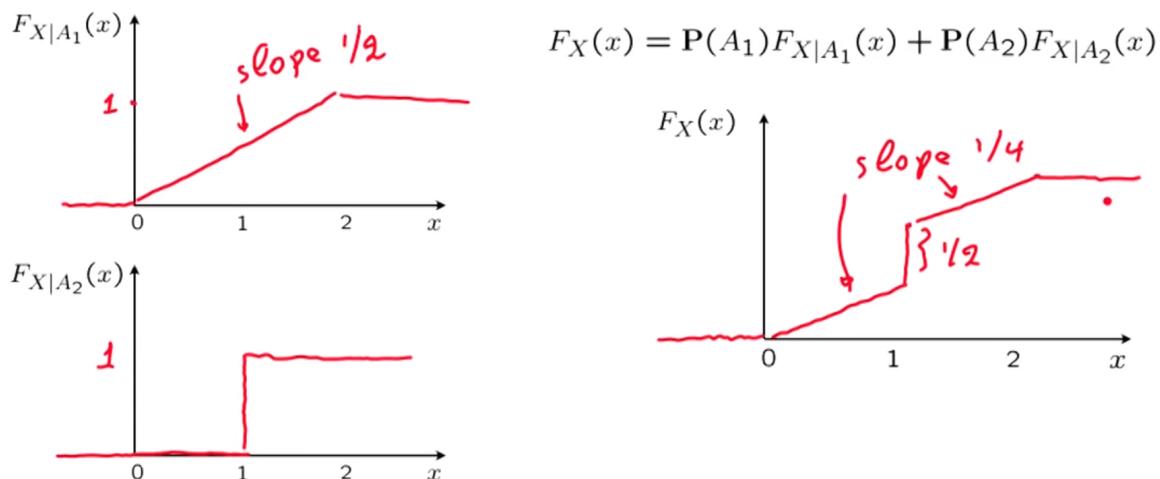
$$\begin{aligned} F_X(x) &= p \cdot P(Y \leq x) + (1-p) P(Z \leq x) \\ &= p F_Y(x) + (1-p) F_Z(x) \end{aligned}$$

$$E[X] = p E[Y] + (1-p) E[Z]$$

## Mixed distributions

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

**A<sub>1</sub>**, **A<sub>2</sub>**



Exercise: A mixed random variable

1/1 point (graded)

A lightbulb is installed. With probability  $1/3$ , it burns out immediately when it is first installed. With probability  $2/3$ , it burns out after an amount of time that is uniformly distributed on  $[0, 3]$ . The expected value of the time until the lightbulb burns out is

✓ Answer: 1

Solution:

The expected value of a uniform on  $[0, 3]$  is  $3/2$ . Using the definition of expectation of mixed random variables, the expected value is  $\frac{1}{3} \cdot 0 + \frac{2}{3} \cdot \frac{3}{2} = 1$ .

### Jointly continuous r.v.'s and joint PDFs

$p_X(x)$	$f_X(x)$
$p_{X,Y}(x, y)$	$f_{X,Y}(x, y)$

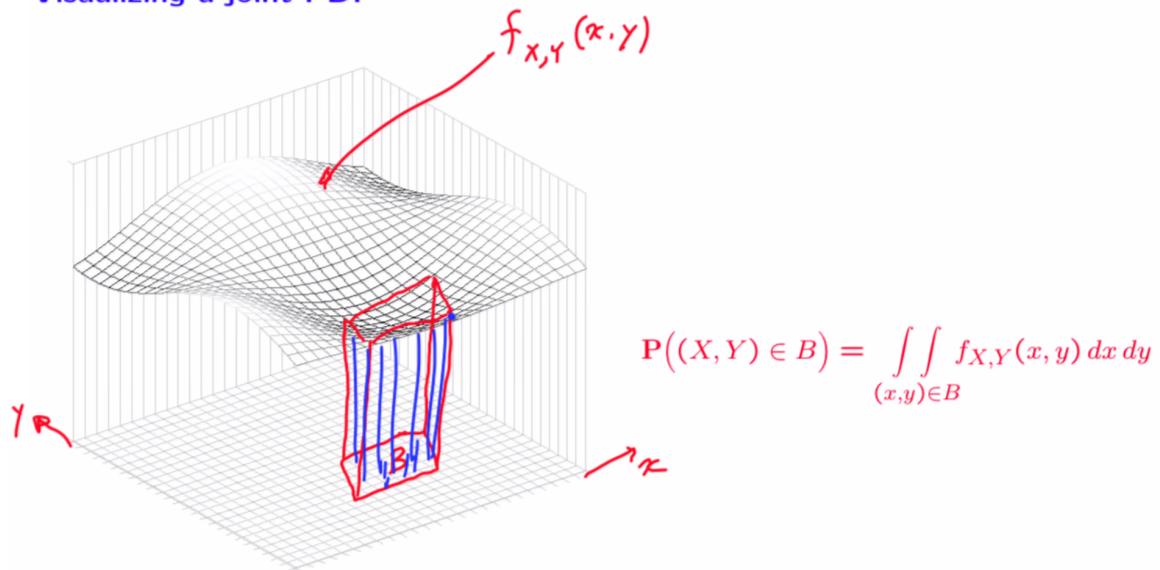
$$p_{X,Y}(x, y) = \mathbf{P}(X = x \text{ and } Y = y) \geq 0 \quad f_{X,Y}(x, y) \geq 0$$

$$\mathbf{P}((X, Y) \in B) = \sum_{(x,y) \in B} p_{X,Y}(x, y) \quad \mathbf{P}((X, Y) \in B) = \int_{(x,y) \in B} f_{X,Y}(x, y) dx dy$$

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

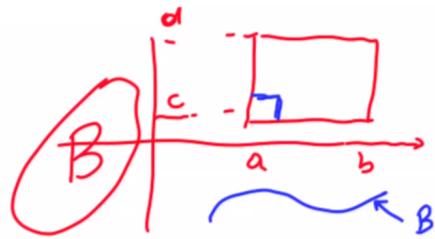
**Definition:** Two random variables are **jointly continuous** if they can be described by a joint PDF

### Visualizing a joint PDF



### On joint PDFs

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

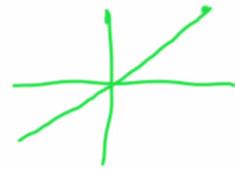


$$\mathbf{P}(a \leq X \leq b, c \leq Y \leq d) = \int_c^d \int_a^b f_{X,Y}(x, y) dx dy$$

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

$Y = X$

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



joint continuity doesn't mean that all r.v.'s are continuous but instead that the probability has to be spread across all their dimensions  
Here  $Y=X$  is not joint continuous because the probability can be defined by one dimension

Exercise: Jointly continuous r.v.'s

2/2 points (graded)

The random variables  $X$  and  $Y$  are continuous. Is this enough information to determine the value of  $\mathbf{P}(X^2 = e^{3Y})$ ?

No ▼ ✓ Answer: No

The random variables  $X$  and  $Y$  are jointly continuous. Is this enough information to determine the value of  $\mathbf{P}(X^2 = e^{3Y})$ ?

Yes ▼ ✓ Answer: Yes

**Solution:**

a) There is no information on the relation between the two random variables. If, for example,  $X = \sqrt{e^{3Y}}$ , the probability is 1, whereas if  $X = \sqrt{e^{3Y}} + 1$ , then the probability is zero.

b) The set of points on the  $x$ - $y$  plane that correspond to the event  $X^2 = e^{3Y}$  is a one-dimensional curve, which has zero area, and therefore zero probability.

## Exercise: From joint PDFs to probabilities

8/8 points (graded)

- a) The probability of the event that  $0 \leq Y \leq X \leq 1$  is of the form  $\int_a^b \left( \int_c^d f_{X,Y}(x, y) dx \right) dy$ . Find the values of  $a, b, c, d$ . Each one of your answers should be one of the following: 0, x, y, or 1.

$a =$   ✓ Answer: 0

0

$b =$   ✓ Answer: 1

1

$c =$   ✓ Answer: y

y

$d =$   ✓ Answer: 1

1

- b) The probability of the event that  $0 \leq Y \leq X \leq 1$  is also of the form  $\int_a^b \left( \int_c^d f_{X,Y}(x, y) dy \right) dx$ . Note the different order of integration as compared to part (a).

Find the values of  $a, b, c, d$ . Each one of your answers should be one of the following: 0, x, y, or 1.

$a =$   ✓ Answer: 0

0

$b =$   ✓ Answer: 1

1

$c =$   ✓ Answer: 0

0

$d =$   ✓ Answer: x

x

**Solution:**

- a) For any given  $y \in [0, 1]$ ,  $x$  ranges from  $y$  to 1, yielding  $\int_0^1 \int_y^1 f_{X,Y}(x, y) dx dy$ .

- b) For any given  $x \in [0, 1]$ ,  $y$  ranges from 0 to  $x$ , yielding  $\int_0^1 \int_0^x f_{X,Y}(x, y) dy dx$ .

### From the joint to the marginals

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

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

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

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

$$F_x(x) = P(X \leq x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f_{X,Y}(s, t) dt ds$$

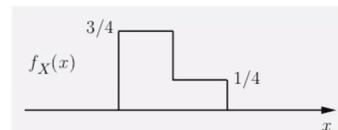
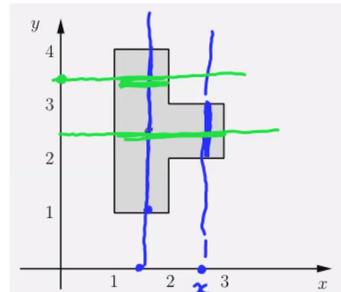
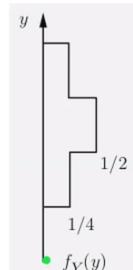
$$f_x(x) = \frac{d F_x(x)}{dx} = [ ]$$

### Uniform joint PDF on a set $S$

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

$$f_{X,Y} = \frac{1}{4}$$

$$\begin{aligned} & \text{area}(A \cap S) \\ & \frac{\text{area}(A \cap S)}{\text{area}(S)} \\ & P(A) \end{aligned}$$



## Exercise: Finding a marginal PDF

1/1 point (graded)

The random variables  $X$  and  $Y$  are described by a uniform joint PDF of the form  $f_{X,Y}(x, y) = 3$  on the set  $\{(x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1, y \leq x^2\}$ .

Then,  $f_X(0.5) =$  0.75

✓ Answer: 0.75

**Solution:**

For any  $x \in [0, 1]$ , and using also the fact that the PDF is zero outside the specified set of  $x$ - $y$  pairs, we have

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy = \int_0^{x^2} 3 dy = 3x^2. \text{ Therefore, } f_X(0.5) = 3/4.$$

## Functions of multiple random variables

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

**Expected value rule:**

$$\mathbf{E}[g(X, Y)] = \sum_x \sum_y g(x, y) p_{X,Y}(x, y) \quad \mathbf{E}[g(X, Y)] = \int \int g(x, y) f_{X,Y}(x, y) dx dy$$

## Linearity of expectations

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

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

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

### Exercise: From joint PDFs to the marginals

2/5 points (graded)

For each one of the following formulas, identify those that are always true. All integrals are meant to be from  $-\infty$  to  $\infty$ .

$$f_{X,Z}(a, b) = \int f_{X,Y,Z}(a', b, c) da'$$

No

✓ Answer: No

$$f_{X,Z}(a, c) = \int f_{X,Y,Z}(a, b, c) db$$

Yes

✓ Answer: Yes

$$f_{X,Z}(a, b) = \int f_{X,Y,Z}(a, b, c) dc$$

Yes

✗ Answer: No

$$f_Y(a) = \int \int \int f_{U,V,X,Y}(a, b, c, s) db dc ds$$

Yes

✗ Answer: No

$$f_Y(a) = \int \int \int f_{U,V,X,Y}(s, c, b, a) db dc ds$$

No

✗ Answer: Yes

#### Solution:

In each case, we need to "integrate out" the arguments associated with random variables that do not appear on the left-hand side. Thus, the correct formulas are:

$$f_{X,Z}(a, c) = \int f_{X,Y,Z}(a, b, c) db$$

and

$$f_Y(a) = \int \int \int f_{U,V,X,Y}(s, c, b, a) db dc ds.$$

## The joint CDF

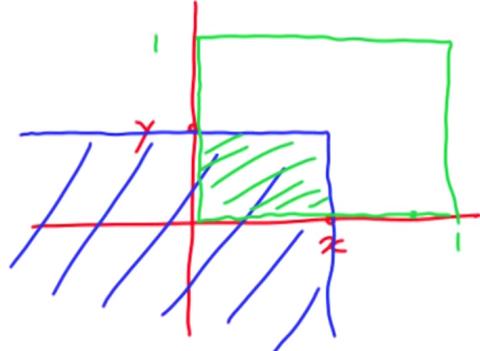
$$F_X(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^x f_X(t) dt \quad f_X(x) = \frac{dF_X}{dx}(x)$$

$$F_{X,Y}(x, y) = \mathbf{P}(X \leq x, Y \leq y) = \int_{-\infty}^y \left[ \int_{-\infty}^x f_{X,Y}(s, t) ds \right] dt$$

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

$$F_{X,Y}(x, y) = xy$$

$$F_{X,Y}(x, y) = 1$$



So by differentiating the CDF with respect to  $y$  and then  $x$  (order doesn't matter) then we get the joint PDF

Exercise: Joint CDFs

3/3 points (graded)

a) Is it always true that if  $x < x'$ , then  $F_{X,Y}(x, y) \leq F_{X,Y}(x', y)$ ?

Answer: Yes

b) Suppose that the random variables  $X$  and  $Y$  are jointly continuous and take values on the unit square, i.e.,  $0 \leq x \leq 1$  and  $0 \leq y \leq 1$ . Is  $F_{X,Y}(x, y) = (x + 2y)^2/9$  a legitimate joint CDF? Hint: Consider  $F_{X,Y}(0, 1)$ .

Answer: No

c) As above, suppose that the random variables  $X$  and  $Y$  are jointly continuous and take values on the unit square, i.e.,  $0 \leq x \leq 1$  and  $0 \leq y \leq 1$ . The joint CDF on that set is of the form  $xy(x + y)/2$ . Find an expression for the joint PDF which is valid for  $(x, y)$  in the unit square. Enter an algebraic function of  $x$  and  $y$  using standard notation.

Answer:  $x+y$

**Solution:**

a) Since  $x < x'$ , the event  $\{X \leq x, Y \leq y\}$  is a subset of the event  $\{X \leq x', Y \leq y\}$ , and therefore  $F_{X,Y}(x, y) = \mathbf{P}(X \leq x, Y \leq y) \leq \mathbf{P}(X \leq x', Y \leq y) = F_{X,Y}(x', y)$ .

b) Since the random variables are nonnegative, we have  $F_{X,Y}(0, 1) = \mathbf{P}(X \leq 0 \text{ and } Y \leq 1) = \mathbf{P}(X = 0 \text{ and } Y \leq 1) \leq \mathbf{P}(X = 0) = 0$ , where the last equality holds because  $X$  is a continuous random variable. But zero is different from  $(0 + 2 \cdot 1)^2/9$ . Therefore, we do not have a legitimate joint CDF.

c) The joint CDF is of the form  $x^2y/2 + y^2x/2$ . The partial derivative with respect to  $x$  is  $xy + y^2/2$ . Taking now the partial derivative with respect to  $y$ , we obtain  $x + y$ .