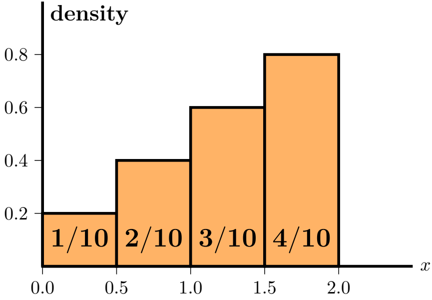
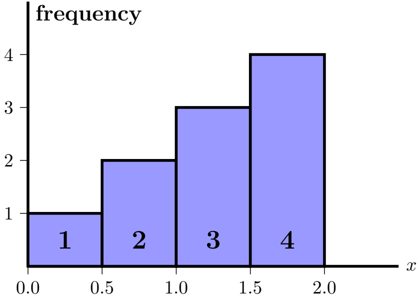
Class 5 (continued): Continuous random variables
18.05 Spring 2022





Announcements/Agenda

Announcements

- None

Agenda

- Gallery of random variables: exponential, uniform, normal
- Histograms

Exponential Random Variables

Parameter: λ (called the rate parameter).

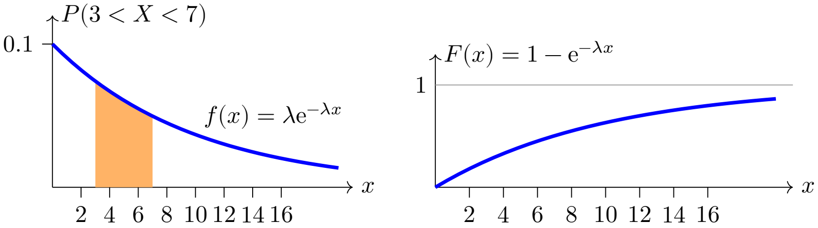
Range: $[0, \infty)$.

Notation: exponential(λ) or exp(λ).

Density: $f(x) = \lambda e^{-\lambda x}$ for $0 \le x$.

Mean: $\dfrac{1}{\lambda}$ Variance $\dfrac{1}{\lambda^2}$

Models: Waiting time



Continuous analogue of geometric distribution!

Board question

I've noticed that taxis drive past 77 Mass. Ave. on the average of once every 10 minutes.

Suppose time spent waiting for a taxi is modeled by an exponential random variable

$$X \sim \text{Exponential}(1/10); \qquad f(x) = \frac{1}{10}\mathrm{e}^{-x/10}$$

(a) Sketch the pdf of this distribution

(b) Shade the region which represents the probability of waiting between 3 and 7 minutes

(c) Compute the probability of waiting between between 3 and 7 minutes for a taxi

(d) Compute and sketch the cdf.

Uniform Random Variables

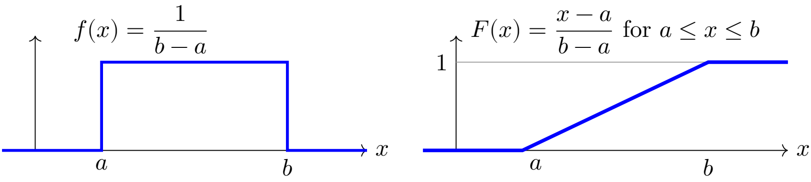
Parameters: a, b (limits).

Range: $[a, b]$.

Notation: $U(a, b)$ or uniform(a, b) or unif(a, b)

Density: $f(x) = \dfrac{1}{b-a}$ (constant)

Mean: $\dfrac{a+b}{2}$ Variance: $\dfrac{(b-a)^2}{12}$



pdf and cdf of uniform(a, b)

Normal Random Variables

Parameters: μ, σ (mean and standard deviation).

Range: $(-\infty, \infty)$.

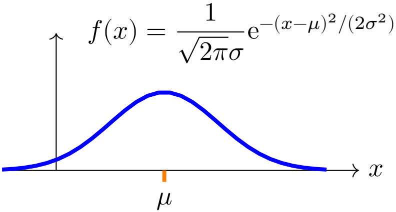
Notation: $\mathrm{N}(\mu, \sigma^2)$

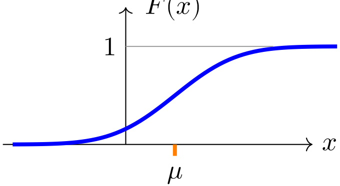
Density: $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-(x-\mu)^2/2\sigma^2}$

Mean: μ Variance: σ^2

Models: Many things: sums of many independent effects.

Standard normal: $\mathrm{N}(0, 1)$, i.e. mean 0, standard deviation 1.





pdf and cdf of $\mathrm{N}(\mu, \sigma^2)$

Table question: Gallery of distributions

Open the Gallery of probability distributions applet at

https://mathlets.org/mathlets/probability-distributions/

(a) For the **standard normal** distribution $\mathsf{N}(0, 1)$ how much probability is within 1 of the mean? Within 2? Within 3?

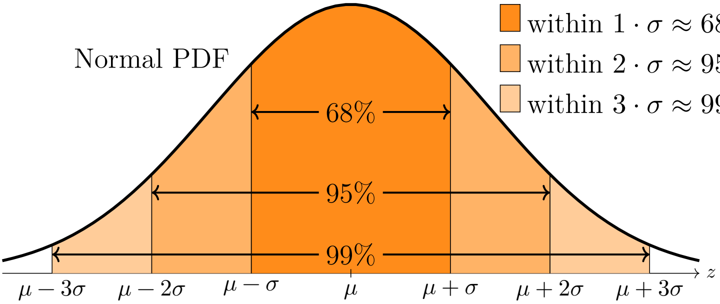
(b) For $\mathsf{N}(0, 3^2)$ how much probability is within σ of the mean? Within 2σ? Within 3σ.

(c) Does changing μ change your answer to problem 2?

(d) Use the applet to find the median of the $\exp(0.5)$ distribution.

(The median is the value of x where half the probability is below x and half above.)

Normal probabilities



Rules of thumb for standard and general normal

$$\begin{aligned}
&P(-1 \le Z \le 1) \approx 0.68, && P(-\sigma \le X - \mu \le \sigma) \approx 0.68\\
&P(-2 \le Z \le 2) \approx 0.95, && P(-2\sigma \le X - \mu \le 2\sigma) \approx 0.95\\
&P(-3 \le Z \le 3) \approx 0.997. && P(-3\sigma \le X - \mu \le 3\sigma) \approx 0.997
\end{aligned}$$

Manipulating random variables

Example. Suppose X has range [0, 2] and cumulative distribution function (cdf) $F_X(x) = x^3/8$ over that range. If $Y = X^2$ find the cdf and pdf for Y.

Manipulating random variables

Example. Suppose X has range [0, 2] and cumulative distribution function (cdf) $F_X(x) = x^3/8$ over that range. If $Y = X^2$ find the cdf and pdf for Y.

Solution: Y has range [0, 4]. To find its cdf you need to remember the definition of the cdf and work carefully through that.

$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(X \le \sqrt{y}) = F_X(\sqrt{y}) = \frac{y^{3/2}}{8}.$$

So the pdf $f_Y(y) = F_Y'(y) = \dfrac{3}{16} y^{1/2}$.

Note. We work with the definition of the cdf as a probability. **Don't guess!** Work systematically.

Board question: Manipulating random variables

(a) Suppose $X \sim$ uniform(0,2). If $Y = 4X$, find the range, pdf and cdf of Y.

(b) Suppose $X \sim$ uniform(0,2). If $Y = X^3$, find the range, pdf and cdf of Y.

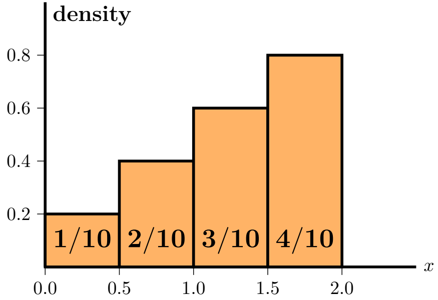
(c) Suppose $Z \sim \text{Norm}(0,1)$ (standard normal). Find the range, pdf and cdf of $Y = 3Z + 2$.

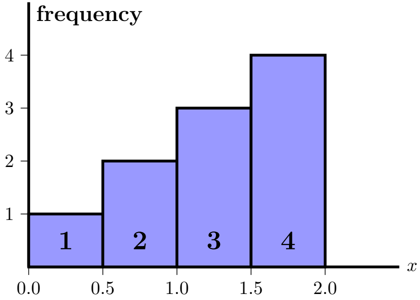
Histograms

Made by 'binning' data.

Frequency: height of bar over bin = number of data points in bin.

Density: area of bar is the fraction of all data points that lie in the bin. So, total area is 1.





Check that the total area of the histogram on the right is 1.

Board question: Histograms

(a) Make both a frequency and density histogram from the data below.

Use bins of width 0.5 starting at 0. The bins should be right closed.

1	1.2	1.3	1.6	1.6
2.1	2.2	2.6	2.7	3.1
3.2	3.4	3.8	3.9	3.9

(b) Same question using unequal width bins with edges 0, 1, 3, 4.

(c) For part (b), why does the density histogram give a more reasonable representation of the data?