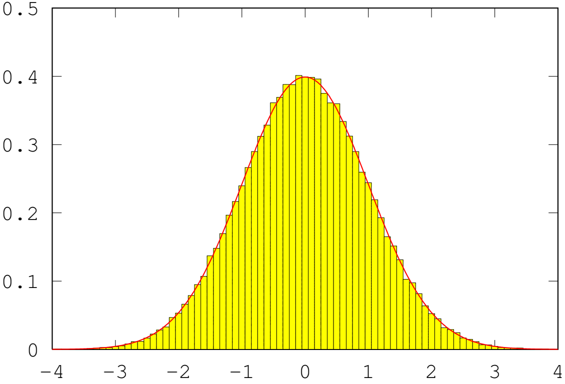
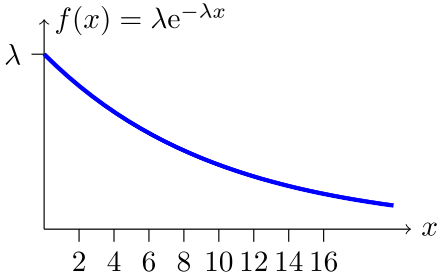
Continuous Expectation and Variance, and the Law of Large Numbers

18.05 Spring 2018



Recall Exponential Random Variables

Parameter: λ (called the rate parameter).
Range: $[0, \infty)$. Expected value: $1/\lambda$
Notation: exponential(λ) or exp(λ).
Density: $f(x) = \lambda e^{-\lambda x}$ for $0 \leq x$.
Models: Waiting time.
large rate $\leftrightarrow$ large λ $\leftrightarrow$ small wait.



Continuous analogue of geometric distribution—memoryless!

Board question

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

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

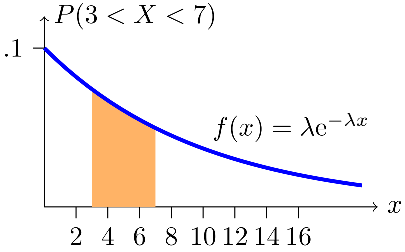
(b) Shade the region which represents the probability of waiting between 3 and 7 minutes

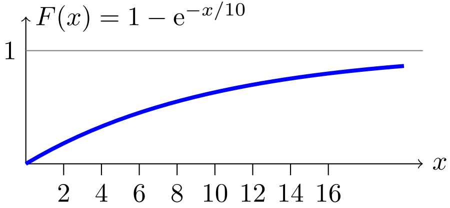
(c) Compute the probability of waiting between between 3 and 7 minutes for a taxi

(d) Compute and sketch the cdf.

Solution

Sketches for (a), (b), (d)





(c)

Expected value

Expected value: measure of location

X continuous with range $[a, b]$ and pdf $f(x)$:

$$E(X) = \int_a^b x f(x)\, dx.$$

X discrete with values $x_1, \ldots, x_n$ and pmf $p(x_i)$:

$$E(X) = \sum_{i=1}^{n} x_i p(x_i).$$

View these as essentially the same formulas.

Variance and standard deviation

Standard deviation: measure of spread, scale

For *any* random variable X with mean μ

$$\mathrm{Var}(X) = E((X-\mu)^2), \qquad \sigma = \sqrt{\mathrm{Var}(X)}$$

X continuous with range $[a, b]$ and pdf $f(x)$:

$$\mathrm{Var}(X) = \int_a^b (x-\mu)^2 f(x)\, dx.$$

X discrete with values $x_1, \ldots, x_n$ and pmf $p(x_i)$:

$$\mathrm{Var}(X) = \sum_{i=1}^{n} (x_i - \mu)^2 p(x_i).$$

View these as essentially the same formulas.

Properties: (the same for discrete and continuous)

1. $E(X + Y) = E(X) + E(Y)$.
2. $E(aX + b) = aE(X) + b$.
3. If X and Y are independent then
   $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.
4. $\text{Var}(aX + b) = a^2\text{Var}(X)$.
5. $\text{Var}(X) = E(X^2) - E(X)^2$.

Board question

The random variable X has range [0,1] and pdf cx^2.

(a) Find c.

(b) Find the mean, variance and standard deviation of X.

(c) Find the median value of X.

(d) Suppose $X_1, \ldots X_{16}$ are independent identically-distributed copies of X. Let $\overline{X}$ be their average. What is the standard deviation of $\overline{X}$?

answer: *See next slides.*

Solution

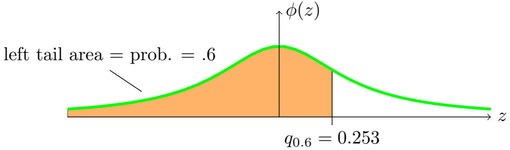
(a) Total probability is 1: $\int_0^1 cx^2\,dx = 1 \;\Rightarrow\; \boxed{c = 3}$.

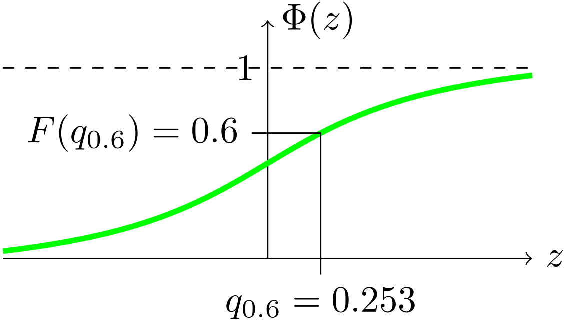
(b) $\mu = \int_0^1 3x^3\,dx = 3/4$.
$\sigma^2 = (\int_0^1 (x - 3/4)^2\, 3x^2\,dx) = \frac{3}{5} - \frac{9}{8} + \frac{9}{16} = \frac{3}{80}$.
$\sigma = \sqrt{3/80} = \frac{1}{4}\sqrt{3/5} \approx .194$

(c) Set $F(q_{0.5}) = 0.5$, solve for $q_{0.5}$: $F(x) = \int_0^x 3u^2\,du = x^3$. Therefore, $F(q_{0.5}) = q_{0.5}^3 = .5$. We get, $\boxed{q_{0.5} = (0.5)^{1/3}}$.

(d) Because they are independent
$\mathrm{Var}(X_1 + \ldots + X_{16}) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_{16}) = 16\mathrm{Var}(X)$.
Thus, $\mathrm{Var}(\overline{X}) = \frac{16\mathrm{Var}(X)}{16^2} = \frac{\mathrm{Var}(X)}{16}$. Finally, $\sigma_{\overline{X}} = \boxed{\frac{\sigma_X}{4} = 0.194/4}$.

Quantiles: a measure of location



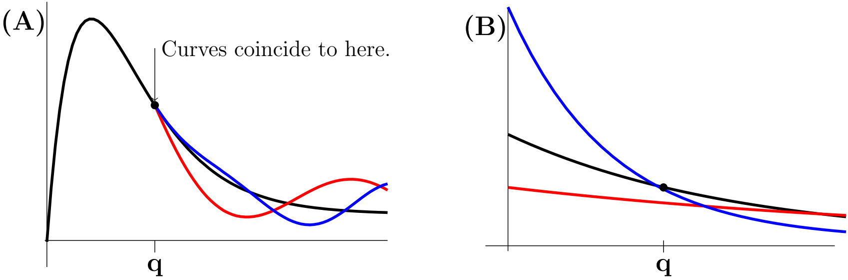


$q_{0.6}$: left tail area $= 0.6 \;\Leftrightarrow\; F(q_{0.6}) = 0.6$

Concept question

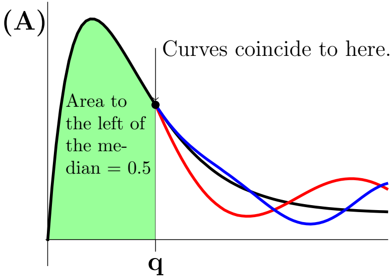
Each of the curves is the density for a given random variable. The median of the black plot is always at q. Which density has the greatest median?

1. Black
2. Red
3. Blue
4. All the same
5. Impossible to tell



answer: *See next frame.*

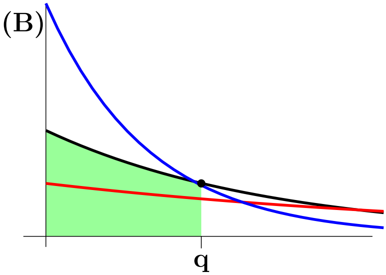
Solution



Plot A: $\boxed{\text{4. All three medians are the same.}}$ Remember that probability is computed as the area under the curve. By definition the median q is the point where the shaded area in Plot A .5. Since all three curves coincide up to q. That is, the shaded area in the figure is represents a probability of .5 for all three densities.

Continued on next slide.

Solution continued



Plot B: $\boxed{\text{2. The red density has the greatest median.}}$ Since q is the median for the black density, the shaded area in Plot B is .5. Therefore the area under the blue curve (up to q) is greater than .5 and that under the red curve is less than .5. This means the median of the blue density is to the left of q (you need less area) and the median of the red density is to the right of q (you need more area).

Law of Large Numbers (LoLN)

- Informally: An average of many measurements is more accurate than a single measurement.

- Formally: Let X_1, X_2, ... be i.i.d. random variables all with mean μ and standard deviation σ.
  Let
  $$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$
  Then for any (small number) a, we have
  $$\lim_{n\to\infty} P(|\overline{X}_n - \mu| < a) = 1.$$

- **No guarantees but:** By choosing n large enough we can make $\overline{X}_n$ as close as we want to μ with probability close to 1.

Concept Question: Desperation

- You have \$100. You need \$1000 by tomorrow morning.
- Your only way to get it is to gamble.
- If you bet \$k, you either win \$k with probability p or lose \$k with probability $1 - p$.

Maximal strategy: Bet as much as you can, up to what you need, each time.
Minimal strategy: Make a small bet, say \$5, each time.

1. If $p = 0.45$, which is the better strategy?
   (a) Maximal (b) Minimal (c) They are the same

Concept Question: Desperation

- You have \$100. You need \$1000 by tomorrow morning.
- Your only way to get it is to gamble.
- If you bet \$k, you either win \$k with probability p or lose \$k with probability $1 - p$.

Maximal strategy: Bet as much as you can, up to what you need, each time.
Minimal strategy: Make a small bet, say \$5, each time.

1. If $p = 0.45$, which is the better strategy?
   (a) Maximal (b) Minimal (c) They are the same

2. If $p = 0.8$, which is the better strategy?
   (a) Maximal (b) Minimal (c) They are the same

answer: *On next slide*

Solution to previous two problems

answer: If $p = 0.45$ use maximal strategy; If $p = 0.8$ use minimal strategy. If you use the minimal strategy the law of large numbers says your average winnings per bet will almost certainly be the expected winnings of one bet. The two tables represent $p = 0.45$ and $p = 0.8$ respectively.

Win	-10	10
p	0.55	0.45

Win	-10	10
p	0.2	0.8

The expected value of a \$5 bet when $p = 0.45$ is -\$0.50 Since on average you will lose \$0.50 per bet you want to avoid making a lot of bets. You go for broke and hope to win big a few times in a row. It's not very likely, but the maximal strategy is your best bet.

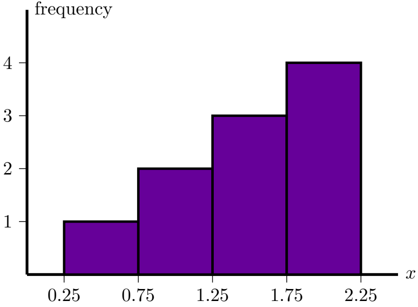
The expected value when $p = 0.8$ is \$3. Since this is positive you'd like to make a lot of bets and let the law of large numbers (practically) guarantee you will win an average of \$6 per bet. So you use the minimal strategy.

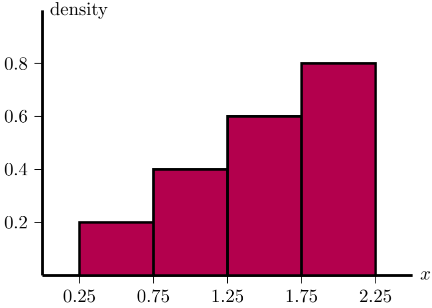
Histograms

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

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

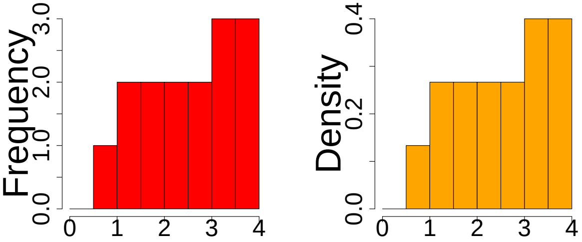
Use bins of width 0.5 starting at 0. The bins should be right closed.

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

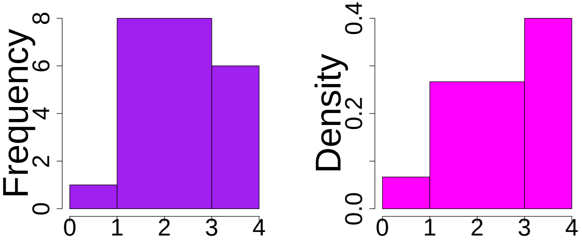
2. Same question using unequal width bins with edges 0, 1, 3, 4.

3. For question 2, why does the density histogram give a more reasonable representation of the data?

Solution



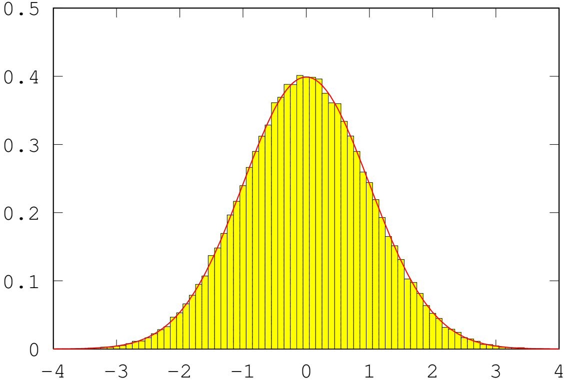
Histograms with equal width bins



Histograms with unequal width bins

LoLN and histograms

LoLN implies density histogram converges to pdf:



Histogram with bin width 0.1 showing 100000 draws from a standard normal distribution. Standard normal pdf is overlaid in red.

Standardization

Random variable X with mean μ and standard deviation σ.

Standardization: $$Y = \frac{X - \mu}{\sigma}.$$

- Y has mean 0 and standard deviation 1.
- Standardizing any normal random variable produces the standard normal.
- If $X \approx$ normal then standardized $X \approx$ stand. normal.
- We reserve Z to mean a standard normal random variable.