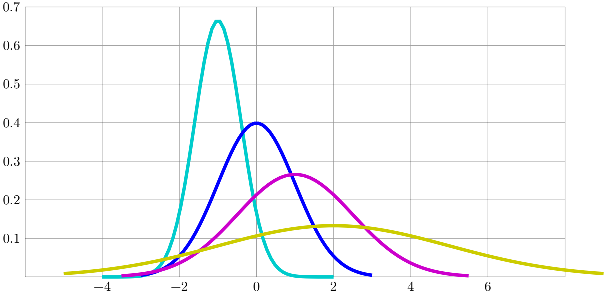
Continuous Expectation and Variance, Quantiles, the Law of Large Numbers
18.05 Spring 2022



Agenda/Announcements

- Friday is a normal class in – we'll finish class 6
- In case of a snow closing, we will cancel class and adjust the pset and schedule.
- R studio graded:
  - Good job!
  - Notice that we are learning how easy simulations are to do
  - -0.25 for stray printouts –this will go up as time goes on
- Gabriel has OH after class.

Agenda

- Expected value
- Variance and standard deviation
- Quantiles (median etc.)
- Histograms
- Law of Large Numbers (LoLN)
- Tomorrow: Central Limit Theorem (CLT)

Expected value

Expected value: measure of location, central tendency

Definition. 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, ..., x_n and pmf $p(x_i)$:

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

Continuous and discrete expectation are essentially the same formulas.

Variance and standard deviation

Standard deviation: measure of spread, scale

Definition. For *any* random variable X with mean μ,

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

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

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

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

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

Continuous and discrete variance are essentially the same formulas.

Properties

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 $f(x) = cx^2$.

(a) Find c.

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

(c) Find the median value of X.

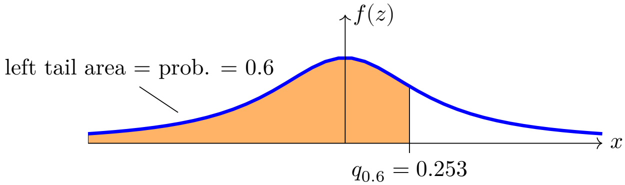
(d) Suppose $X_1, \ldots X_{16}$ are independent identically-distributed copies of X. Let $\overline{X}$ be their average. What is the standard deviation of $\overline{X}$?

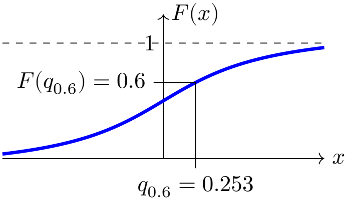
(e) Suppose $Y = X^4$. Compute $E[Y]$

(f) Find the pdf of Y.

Quantiles: measure of location

Example. The 0.6 quantile $q_{0.6}$ is the x-value, with $P(X \leq q_{0.6}) = F(q_{0.6}) = 0.6$.





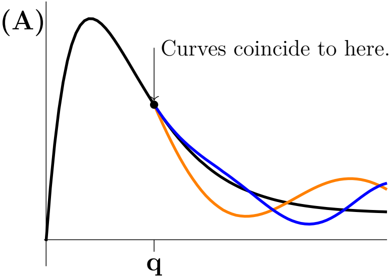
$q_{0.6}$: left tail area $= 0.6 \;\Leftrightarrow\; F(q_{0.6}) = 0.6 \;\Leftrightarrow\; q_{0.6} = F^{-1}(0.6)$

In R: `qnorm, qbinom, qexp etc.` The posted problems for today include one on using R for quantiles.

Concept questions: Greatest median 1

Each of the curves is the density for a random variable. Where there is just one curve they overlap.

The median of the black plot is at q. Which density has the greatest median?

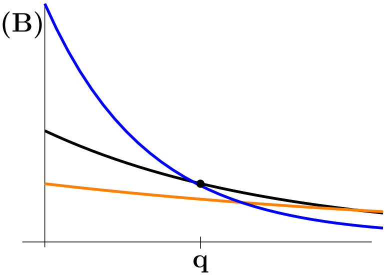


1. Black
2. Orange
3. Blue
4. All the same
5. Impossible to tell

Concept questions: Greatest median 2

Each of the curves is the density for a random variable. Where there is just one curve they overlap.

The median of the black plot is at q. Which density has the greatest median?



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

Histograms

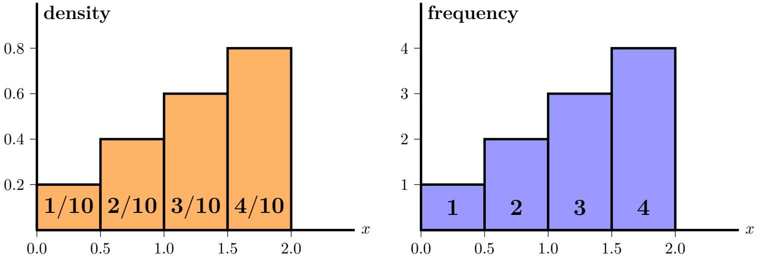
Made by 'binning' data.

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

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

Example: equal bin widths

Consider data 0.5, 0.9, 1.0, 1.3, 1.4, 1.5, 1.8, 1.8, 2.0, 2.0.
Make frequency and density histograms with bin width 0.5 starting at 0.0.
Check that the total area of the histogram on the left is 1.

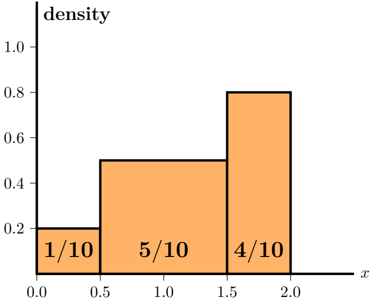


- Values on bin boundaries, e.g. 0.5, 1, 1.5, 2 go to left-hand bin.
- Bins are **right-closed**, e.g first bin is $(0, 0.5]$.
- With equal bin widths the histograms look the same. Only vertical scale is changed.

Example: unequal bin widths

Repeat the example with *unequal* bin widths. Put the bin bounds at 0.0, 0.5, 1.5, 2.0.

Solution: With unequal bin widths the density and frequency histograms look different



Don't be fooled! These are based on the same data.

- If using unequal bin widths, always use a density histogram.
- R will complain if you don't
- Density histogram looks similar to previous histograms.
- Frequency histogram is different and misleading

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?

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.

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

  Then for any (small number) a, we have

$$\lim_{n \to \infty} P(|\overline{X}_n - \mu| < a) = 1.$$

- By choosing n large enough we can, with probability close to 1, make $\overline{X}_n$ as close as we want to μ.

  $\overline{X}_n$ is random, so there may be a small probability that it is far from μ.

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.

(a) 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 $1 - p$.

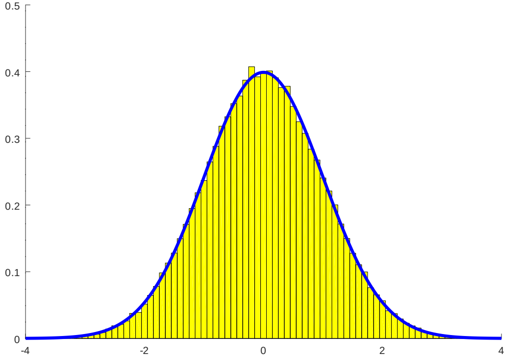
Maximal strategy: Bet as much as you can, up to what you need, each time.
Minimal strategy: Make a small bet, say \$5, each time.

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

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

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 blue.