

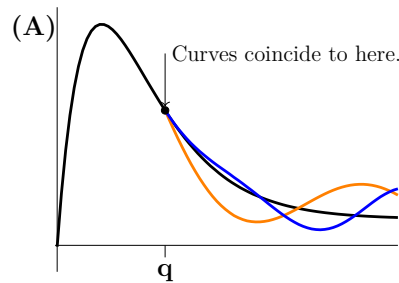
Class 6 in-class problems, 18.05, Spring 2022

Concept questions

Concept question 1. Greatest median 1

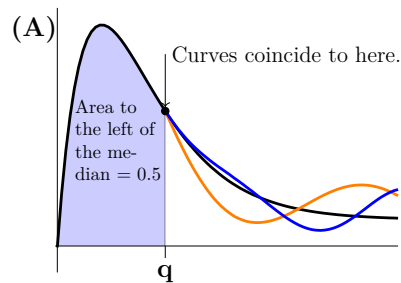
Each of the curves is the density for a random variable. Where there is just one curve they overlap.

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

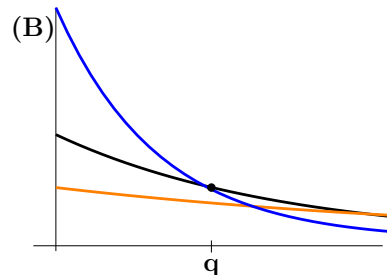
4. All three medians are the same. Remember that probability is computed as the area under the curve. By definition the median q is the point where the shaded area in Plot A is 0.5. Since all three curves coincide up to q . That is, the shaded area in the figure is represents a probability of 0.5 for all three densities. Thus, it is the median for all three densities.



Concept question 2. Greatest median 2

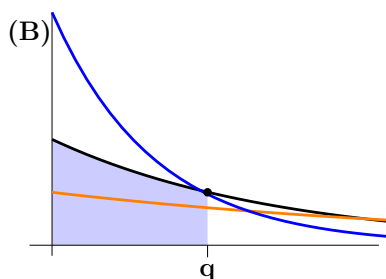
Each of the curves is the density for a random variable. Where there is just one curve they overlap.

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

Solution: 2. The orange density has the greatest median. Since q is the median for the black density, the shaded area in Plot B is 0.5. Therefore the area under the blue curve (up to q) is greater than 0.5 and that under the orange curve is less than 0.5. This means the median of the blue density is to the left of q (you need less area) and the median of the orange density is to the right of q (you need more area).



Concept question 3. 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

Solution: 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	-5	5
p	0.55	0.45

Win	-5	5
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 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 of a \$5 bet 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 \$3 per bet. So you use the minimal strategy.

Board questions**Problem 1.**

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, \dots, X_{16} are independent identically-distributed copies of X . Let \bar{X} be their average. What is the standard deviation of \bar{X} ?

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

(f) Find the pdf of Y .

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

(b) **Solution:** $\mu = \int_0^1 xf(x) dx = \int_0^1 3x^3 dx = 3/4$.

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

$$\sigma = \sqrt{3/80} = \frac{1}{4}\sqrt{3/5} \approx 0.194$$

(c) **Solution:** By definition: $F(q_{0.5}) = 0.5$, so we solve for $q_{0.5}$.

$$F(x) = \int_0^x 3u^2 du = x^3. \text{ Therefore, } F(q_{0.5}) = q_{0.5}^3 = 0.5.$$

$$\text{We get, } \boxed{q_{0.5} = (0.5)^{1/3}}.$$

(d) **Solution:** Because they are independent

$$\text{Var}(X_1 + \dots + X_{16}) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_{16}) = 16\text{Var}(X).$$

$$\text{Thus, } \text{Var}(\bar{X}) = \frac{16\text{Var}(X)}{16^2} = \frac{\text{Var}(X)}{16}.$$

$$\text{Finally, } \sigma_{\bar{X}} = \boxed{\frac{\sigma_X}{4} = 0.194/4}.$$

(e) **Solution:** $E[Y] = \int_0^1 x^4 cx^2 dx = \frac{cx^7}{7} \Big|_0^1 = \frac{c}{7} = \frac{3}{7}$.

(f) **Method 1, use the cdf:**

$$F_Y(y) = P(X^4 < y) = P(X < y^{1/4}) = F_X(y^{1/4}) = y^{3/4}.$$

$$\text{Now differentiate. } f_Y(y) = F'_Y(y) = \boxed{\frac{3}{4}y^{-1/4}}.$$

Method 2, use the pdf:

We have $y = x^4$. So,

$$dy = 4x^3 dx \Rightarrow \frac{dy}{4y^{3/4}} = dx.$$

This implies

$$f_X(x) dx = f_X(y^{1/4}) \frac{dy}{4y^{3/4}} = \frac{3y^{2/4} dy}{4y^{3/4}} = \frac{3}{4y^{1/4}} dy.$$

Therefore. $f_Y(y) = \frac{3}{4y^{1/4}}$

Problem 2.

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

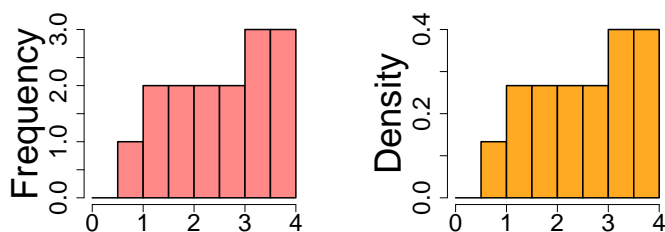
Use bins of width 0.5 starting at 0. The bins should be right closed.

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

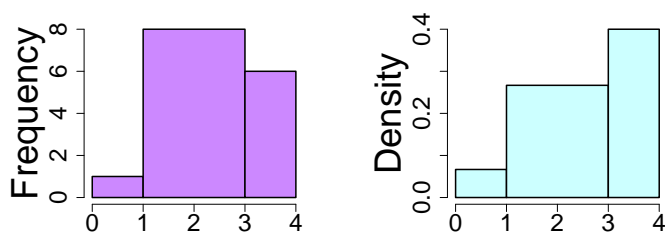
(b) Same question using unequal width bins with edges 0, 1, 3, 4.

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

Solution:



Histograms with equal width bins



Histograms with unequal width bins

Extra problems

Extra 1. Quantiles using R

(a) Let Z be a standard normal variable ($Z \sim N(0,1)$). Use the R function `qnorm` to find the 0.25, 0.5, 0.75 quantiles for Z .

Use the R help to learn about `qnorm`.

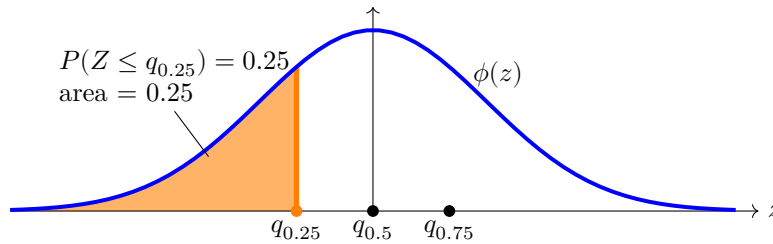
(b) Graph the pdf of the standard normal distribution. Place the quantiles from part (a) on your graph. Also, indicate on the graph the probabilities connected to the quantiles.

(c) The R function `pnorm` is the CDF for the normal distribution. In its simplest form `pnorm(z)` is the CDF for the standard normal distribution.

Apply `pnorm` to your quantiles from part (a). Check that the resulting probabilities are exactly what you expect.

Solution: (a) Using `qnorm(0.25)`, `qnorm(0.5)`, `qnorm(0.75)` we find the quantiles $q_{0.25} \approx -0.6744898$, $q_{0.5} = 0.0$, $q_{0.75} = 0.6744898$.

(b) $q_{0.25}$ is the value of z such that $P(Z \leq q_{0.25}) = 0.25$. This is shown in the figure below. Similar pictures hold for $q_{0.5}$ and $q_{0.75}$.



(c) We find `pnorm(q_0.25) = pnorm(-0.6744898) = 0.25`, `pnorm(q_0.5) = pnorm(0) = 0.5`, `pnorm(q_0.75) = pnorm(0.6744898) = 0.75`.

By definition, $q_{0.25}$ such that $P(Z \leq q_{0.25}) = 0.25$. So the values are just what we expect.

MIT OpenCourseWare

<https://ocw.mit.edu>

18.05 Introduction to Probability and Statistics

Spring 2022

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.