Math 3339

Section 27204
MWF 10-11:00am AAAud 2

Bekki George
bekki@math.uh.edu
639 PGH

Office Hours:
M & Th noon – 1:00 pm & T 1:00 – 2:00 pm
and by appointment

Quick review of continuous distributions (so far): Discrete: $\sum_i f(x_i) = 1$

- A density function is a nonnegative function f defined on the set of real numbers $\mathbb{R}$ such that $\int_{-\infty}^{\infty} f(x)\,dx = 1$
- If *X* is a random variable with probability density function *f*, then the *cumulative distribution function* (abbreviated c.d.f) is the function given by

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(w)\,dw \qquad F'(x) = f(x)$$

$P(X < x)$

- The expected value of *f* is: $E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$
- Generally - $E[h(X)] = \int_{-\infty}^{\infty} h(x) \cdot f(x)\,dx$ (so, $E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx$)
- Variance: $Var[X] = E[X^2] - E[X]^2$

Suppose a rv has pdf:

$$f(x) = \begin{cases} \dfrac{5x^4}{32} & 0 \le x \le k \quad (k = 2) \\ 0 & otherwise \end{cases}$$

$$\int_0^k \frac{5}{32} x^4 \, dx = 1$$

$$\frac{1}{32} x^5 \Big|_0^k = 1$$

$$\frac{k^5}{32} - 0 = 1$$

$$\boxed{k = 2}$$

Find k so that this is a valid pdf

Find: $P(x \le 1)$

$$\int_0^1 \frac{5x^4}{32} \, dx = \frac{x^5}{32} \Big|_0^1 = \frac{1}{32}$$

$P(x \ge 1)$

$$\int_1^2 \frac{5x^4}{32} \, dx$$

$$\text{OR } 1 - P(X \le 1) = \frac{31}{32}$$

Find c so that $P(x \le c) = 0.85$

$$\int_0^c \frac{5x^4}{32} \, dx = .85$$

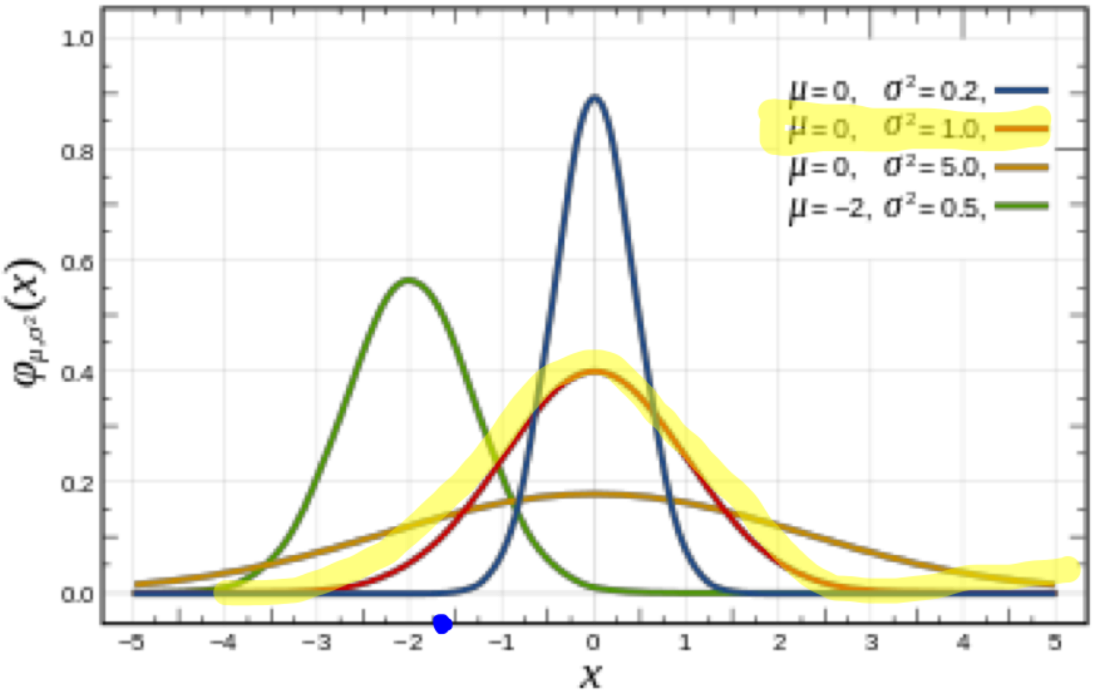
$$\frac{1}{32} x^5 \Big|_0^c = .85$$

$$\frac{c^5}{32} = .85$$

$$c^5 = 27.2$$

$$c \approx 1.936$$

Normal Distributions continued:



$Z \sim \text{Normal}(0,1)$

How many st. dev. we are from mean

Popper 12

$\mu = 0 \quad \sigma = 1$

1. A student takes a standardized exam. The grader reports the student's standardized score (z-score) as -1.8. This indicates:

   a. The student scored lower than the average. ✓

   b. The student scored less than one standard deviation from the average.

   more than

   c. A mistake has been made in calculating the score, since a standard score can never be negative.

   d. Both a and b, but not c.

2. Suppose X is an exponential distribution with rate $\lambda = 4$. What is $E[X]$? $= \mu = 1/\lambda$

a. 4
b. 0.25
c. 16
d. 0.0625
e. none of these

$\mu = \alpha\beta \qquad \sigma^2 = \alpha\beta^2 = \mu \cdot \beta$

3. Suppose we are given a gamma distribution whose mean is 8 and standard deviation is 4. Find α and β

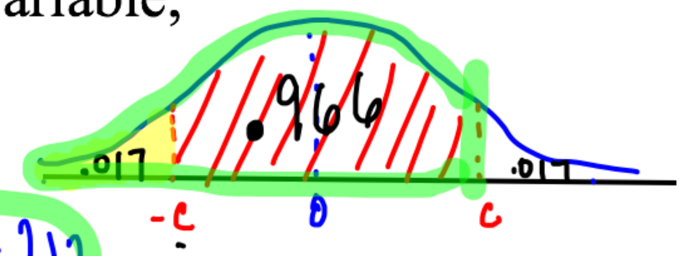
$\mu = 8 \qquad \sigma^2 = 16$

$16 = 8 \cdot \beta$

a. $\alpha = 2$ and $\beta = 2$
b. $\alpha = 2$ and $\beta = 4$
c. $\alpha = 4$ and $\beta = 2$
d. $\alpha = 4$ and $\beta = 4$
e. none of these

$P(Z \le c) = \text{qnorm(value)}$

Example: If Z is the standard normal random variable,



Find a value of c so that $P(-c \le Z \le c) = 0.966$

$c = 2.12$

$\boxed{\text{qnorm}(.966 + .017) = 2.12}$

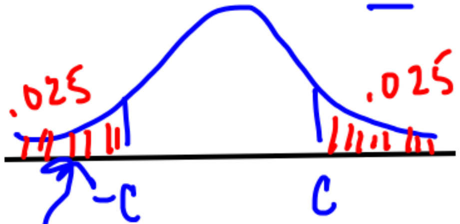
$1 - .966 = .034$

$\frac{1}{2}(.034) = .017$

$\text{qnorm}(.017) = -2.12$

$= -c$

Find a value of c so that $P(|Z| > c) = 0.05$

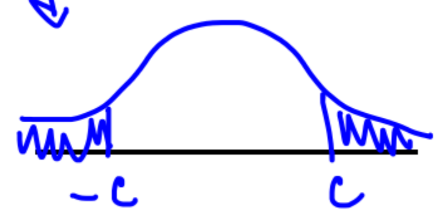


$\text{qnorm}(.025) = -1.96 = -c$

$\boxed{c = 1.96}$

☆ $|Z| < c \iff -c < Z < c$

$|Z| > c \iff Z < -c$ or $Z > c$



$\mu = 0 \quad \sigma = 1$

The table for the Standard Normal Random Variable Z ONLY applies to the standard normal random variable. What do we do if our random variable is not standard?

We may need to make it so. We *transform* it into a standard normal rv.

Suppose X has mean μ, how would we transform X so that the new rv has mean 0?

$X - \mu_X$

Suppose X has variance σ^2, how would we transform X so that it has variance 1?

$$\frac{X}{\sigma_X^2}$$

some books use $N(\mu, \sigma)$

Here's the bottom line: If $X \sim N(\mu, \sigma^2)$, then $\boxed{\dfrac{X-\mu}{\sigma} \sim Z.}$ ☆

Standardizing a Normal Random Variable:

If $X \sim N(\mu, \sigma^2)$, then $P(a < X < b) = P\left(\dfrac{a-\mu}{\sigma} < Z < \dfrac{b-\mu}{\sigma}\right)$, where $Z = \dfrac{X-\mu}{\sigma}$.

Example: Jill scores 710 on the mathematics part of the SAT. The distribution of SAT scores in a reference population is N(500, 100). Jack takes the ACT mathematics test and scores 29. ACT scores follow the distribution N(18, 6). Find the standardized scores for each. Who do you think performed better on their test?

$$\text{Jill}(z) = \frac{710-500}{100} = 2.1$$

$$\text{Jack}(z) = \frac{29-18}{6} = 1.83$$

Example: Say that the time X required to assemble an item is a normally distributed random variable with mean $\mu = 15.8$ minutes and standard deviation $\sigma = 2.4$ minutes.
Determine the probability that the next item will take more than 17 minutes to assemble.

$$\underbrace{P(X > 17)}_{\downarrow \; 1 - \text{pnorm}(17, 15.8, 2.4)} = P\left(Z > \frac{17-15.8}{2.4}\right) = P(Z > .5)$$

$$= 1 - P(Z < .5)$$

$$= 1 - \text{pnorm}(.5)$$

```
> 1-pnorm(17,15.8,2.4)
[1] 0.3085375
> 1-pnorm(.5)
[1] 0.3085375
```

norm.

Suppose an applicant needs to score better than 77% of all GRE test takers to get accepted into this university.

$\mu = 151.3 \quad \sigma = 8.7$

a) What is the minimum score required to meet this criteria?

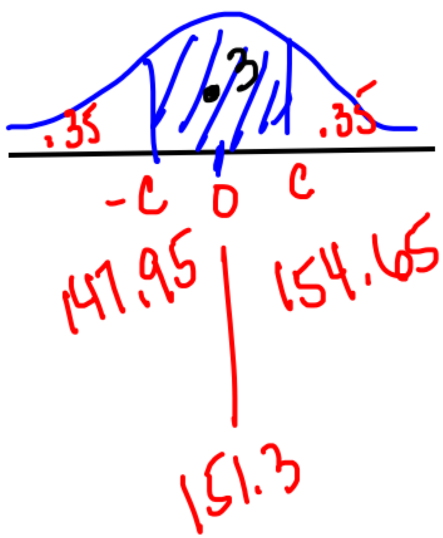
.7388
\# of s.d. above mean

qnorm(.77, 151.3, 8.7

```
> qnorm(.77,151.3,8.7)
[1] 157.728
```

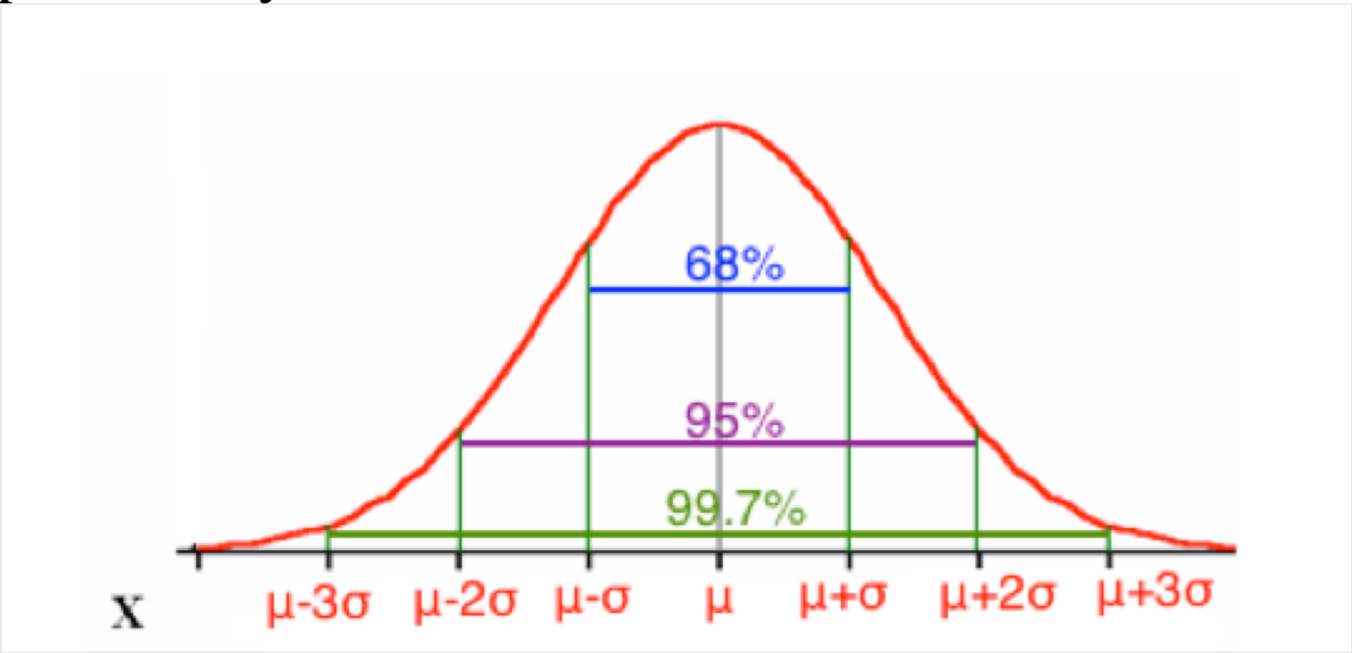
b) What are the cutoff scores that would capture the middle 30% of applicants?

(147.95, 154.65)



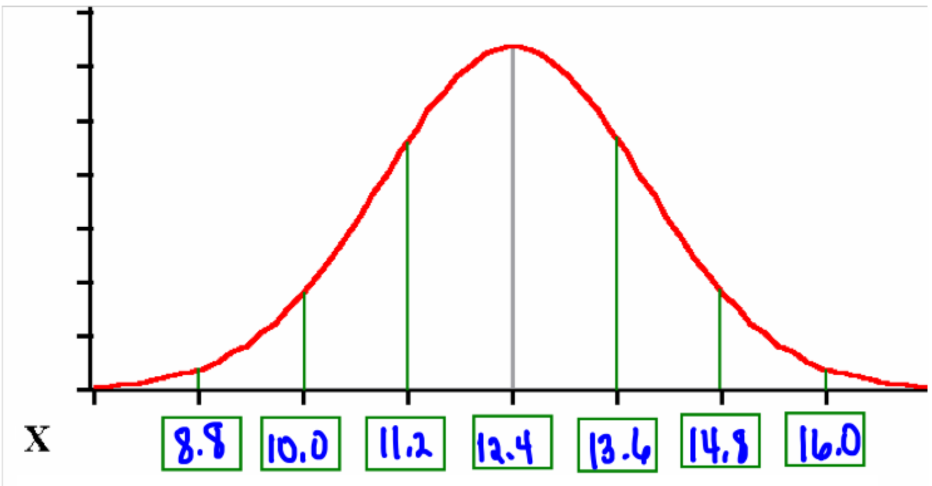
Common and recurring z – values:

1. Approximately 68% of normally distributed observations lie within 1 standard deviation of the mean.
2. Approximately 95% lie within 2 standard deviations of the mean.
3. Approximately 99.74% lie within 3 standard deviations of the mean.



Example

A trucking firm determines that its fleet of trucks averages a mean of 12.4 miles per gallon with a standard deviation of 1.2 miles per gallon on cross-county hauls.



Example: A lunch stand in the business district has a mean daily gross income of \$420 with a standard deviation of \$50. Assume that daily gross incomes are normally distributed.

If a randomly selected day has a gross income of \$500, then how many standard deviations away from the mean is that day's gross income?

$$\frac{500-420}{50} = 1.6 \leftarrow \text{zscore} = \#\text{ of s.d. from mean}$$

What percent of daily gross incomes are between \$300 and \$500?

$$P(300 < X < 500) = \text{pnorm}(500,420,50) - \text{pnorm}(300,420,50)$$
$$= .937$$

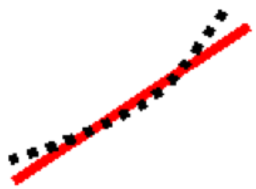
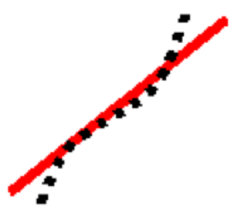
What percent of daily gross incomes are at least \$600?

$$P(X \geq 600) = 1 - \text{pnorm}(600,420,50) = .00016$$

How often (what percent of the time) can the lunch stand expect to make less than \$400?

$$P(X < 400) = \text{pnorm}(400, 420, 50) = .3446$$

Normal Probability Plots

If your chart looks like this:	It indicates that your distribution has:
	Right Skew - If the plotted points appear to bend up and to the left of the normal line that indicates a long tail to the right.
	Left Skew - If the plotted points bend down and to the right of the normal line that indicates a long tail to the left.
	Short Tails - An S shaped-curve indicates shorter than normal tails, i.e. less variance than expected.
	Long Tails - A curve which starts below the normal line, bends to follow it, and ends above it indicates long tails. That is, you are seeing more variance than you would expect in a normal distribution.

from http://www.skymark.com/resources/tools/normal_test_plot.asp

R commands:
qqnorm(*variable*) and qqline(*variable*)
To determine normality of data, look at both the normal probability plots as well as the histogram for the data.

Popper

4. Which of the following are true statements?

   I. The area under a density curve is always 1, regardless of mean and standard deviation.

   II. The mean is equal to the median in a normal distribution.

   III. The interquartile range for any normal curve extends from $\mu - 1\sigma$ to $\mu + 1\sigma$.

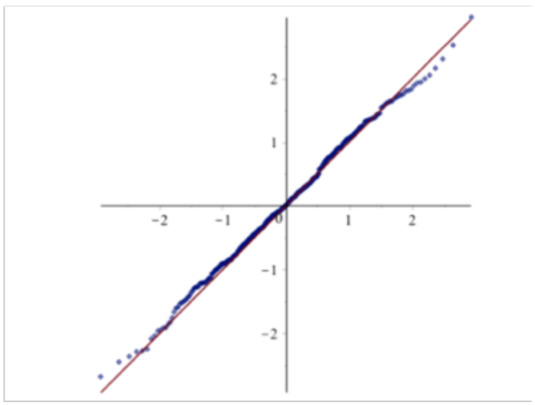
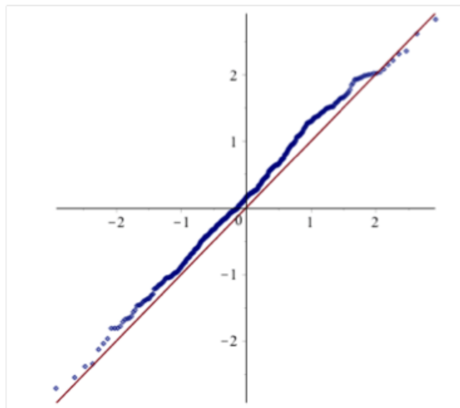
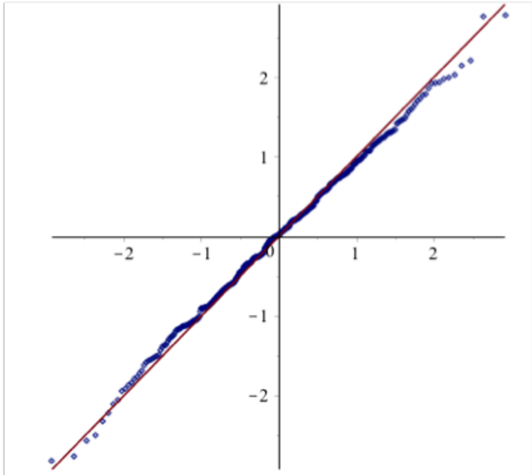
   A. I and II

   B. I and III

   C. II and III

   D. I, II and III

5. Which of the probability plots indicates a good model for a normal distribution?

a)  b)  c) 

d) they all would be good models for a normal distribution