Math 10A with Professor Stankova
Worksheet, Discussion #34; Wednesday, 11/15/2017
GSI name: Roy Zhao

Logistic Regression

Example

1. Given a logistic equation $F(t) = \dfrac{1}{1 + Ce^{-kt}}$, what line do you get after logistic regression?

> **Solution:** In order to do a logistic regression, instead of plotting F vs t, we graph $\ln(F/(1-F))$ vs t. The reason to do this is because
>
> $$\frac{F}{1-F} = \frac{1/(1+Ce^{-kt})}{1 - 1/(1+Ce^{-kt})} = \frac{1}{Ce^{-kt}} = \frac{1}{C}e^{kt}.$$
>
> So taking the log gives us $\ln(F/(1-F)) = \ln(1/C) + \ln(e^{kt}) = kt + \ln(1/C)$. Therefore, the slope is k and the y intercept is $\ln(1/C)$.

Problems

2. **TRUE** False We can only use logistic regression if the data values have y values that lie between 0 and 1 exclusive.

> **Solution:** In order to transform the data, we need to take $\ln(F/(1-F))$ and we can only do this if F is between 0 and 1.

3. **TRUE** False We commonly use logistic regression to model probability of success/failure.

> **Solution:** An example of something we would use logistic regression for is the probability of passing the class versus how many hours you studied. This is because at you go to ∞ hours, your probability of passing goes to 1 and going towards 0 hours goes to a probability of 0. So, this is similar to a logistic function and we could try to model this as a logistic function using logistic regression.

Z-Scores

Example

4. Let $f(x) = \frac{1}{2\sqrt{2\pi}} e^{-(x-5)^2/8}$ be a PDF. Calculate the probability $P(3 \leq X \leq 7)$.

Solution: The formula for a normal distribution is $\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$, and so we see that $\mu = 5$ and $\sigma = 2$. We split up the probability $P(3 \leq X \leq 7)$ at the mean to get $P(3 \leq X \leq 7) = P(3 \leq X \leq 5) + P(5 \leq X \leq 7)$. We use z scores. In order to calculate z scores, we take $\frac{|a-\mu|}{\sigma}$ where a is the value you want to take the z score of. So, the z score of 3 is $\frac{|3-5|}{2} = 1$ and the z score of 7 is $\frac{|7-5|}{2} = 1$. So the probability is

$$z(1) + z(1) = 0.3413 + 0.3413 = 0.6826.$$

Problems

5. **TRUE** False We can only use the z score to calculate probabilities of normal distributions (bell curves).

Solution: The table only applies for probability of normal distributions.

6. True **FALSE** The normal distribution with positive mean can only take on positive values. ($P(X \leq 0) = 0$)

Solution: The normal distribution can take on any real value.

7. Let f be normally distributed with mean 1 and standard deviation 4. Calculate the probability $P(X \geq 3)$.

Solution: In order to calculate this probability, we need the probability area to touch the median so we have $P(X \geq 3) = P(X \geq 1) - P(1 \leq X \leq 3)$. The first probability is $\frac{1}{2}$ and the second has the z score $\frac{|3-1|}{4} = \frac{1}{2}$. So the answer is $0.5 - z(0.5)$.

8. Let f be normally distributed with mean -2 and standard deviation 4. Calculate the probability $P(-1 \leq X \leq 1)$.

Solution: We have $P(-1 \le X \le 1) = P(-2 \le X \le 1) - P(-2 \le X \le -1)$ and we calculate the z scores. The first is $\frac{|1-(-2)|}{4} = \frac{3}{4}$ and the second is $\frac{|-1-(-2)|}{4} = \frac{1}{4}$. Thus, the probability is $z(0.75) - z(0.25)$.

9. Let f be normally distributed with mean -2 and standard deviation 4. Calculate the probability $P(-3 \le X \le 1)$.

Solution: We have $P(-3 \le X \le 1) = P(-3 \le X \le -2) + P(-2 \le X \le 1)$ and we calculate the z scores. The first is $\frac{|-3-(-2)|}{4} = \frac{1}{4}$ and the second is $\frac{|1-(-2)|}{4} = \frac{3}{4}$. Thus, the probability is $z(0.25) + z(0.75)$.

10. Let f be normally distributed with mean 5 and standard deviation 2. Calculate the probability $P(X \le 3)$.

Solution: We have $P(X \le 3) = P(X \le 5) - P(3 \le X \le 5)$ and we calculate the z scores. The first is just $\frac{1}{2}$ and the second is $\frac{|3-5|}{2} = 1$. Thus, the probability is $0.5 - z(1)$.

11. Let f be normally distributed with mean 3 and standard deviation 5. Calculate the probability $P(X \ge 0)$.

Solution: We have $P(X \ge 0) = P(0 \le X \le 3) + P(3 \le X)$ and we calculate the z scores. The second is just $\frac{1}{2}$ and the first is $\frac{|0-3|}{5} = 0.6$. Thus, the probability is $0.5 + z(0.6)$.

12. Let f be normally distributed with mean 2 and standard deviation 1. Calculate the probability $P(X \le 0)$.

Solution: We have $P(X \le 0) = P(X \le 2) - P(0 \le X \le 2)$ and we calculate the z scores. The first is just $\frac{1}{2}$ and the second is $\frac{|0-2|}{1} = 2$. Thus, the probability is $0.5 - z(2)$.

13. Let f be normally distributed with mean 0 and standard deviation 5. Calculate the probability $P(-2 \le X \le -1)$.

Solution: We have $P(-2 \leq X \leq -1) = P(-2 \leq X \leq 0) - P(-1 \leq X \leq 0)$ and we calculate the z scores. The first is $\frac{|-2-0|}{5} = \frac{2}{5}$ and the second is $\frac{|-1-0|}{5} = \frac{1}{5}$. Thus, the probability is $z(2/5) - z(1/5)$.