

Assignment 2

This assignment will give some practice with interpreting studies and computing and interpreting descriptive statistics

1. (Taken from Bland and Peacock) Gardosi compared giving birth in the conventional recumbent manner versus upright squatting using a birth cushion. All participants were full term mothers with a singleton pregnancy and no known antenatal risk factors for complicated delivery. Women were initially randomized to the recumbent versus upright treatment with 209 assigned to the recumbent treatment versus 218 randomized to the upright treatment. However, 22 women in the recumbent group spontaneously requested and were allowed to have a cushion and deliver in the upright position and 39 assigned to the upright position were comfortable lying down and delivered in the recumbent position.

The authors compared the two groups by **intent to treat**, that is, by their original randomized assignment. The table below shows a primary outcome, whether the delivery required forceps or not, by intent to treat

	<u>Upright</u>	<u>Recumbent</u>
Num forceps Deliveries (pct)	19/218=8.7%	34/209=16.3% (plus 2 others required C section)

1a. This result implies that Upright has fewer forceps deliveries. In a few sentences, briefly discuss the differences between intent to treat and on-treatment (as treated) analysis. What is the directionality in the biases they have, if any? For example - if no one had changed groups, would we expect to see a bigger or smaller difference between upright and recumbent? Does treatment group “crossover” cause the groups to appear to be more similar or more distinct?

1b. Can this study be blinded and does this matter? Brief answer.

2. The table below shows the mean and SD for duration of the second stage of labor from this study

	second stage of labor duration in minutes	
	<u>Upright</u>	<u>Recumbent</u>
Median	31	45
Mean	39	50
SD	26	29
IQR	23-50	28-67 (Interquartile range=Q3-Q1)

Do the results imply that the distribution of duration time may be well modeled by a symmetric Gaussian (normal bell curve)? Explain why or why not or why one cannot determine this based on the table. Is a mean or median the best choice as a measure of “typical” duration?

3. The data below show the mean and standard deviation (SD) for weight loss (or gain) in kg in two different diet groups after 12 months of dieting. (Gardner 2007)

Weight loss in kg		
	Zone	Atkins
Mean	1.6	4.7
SD	4.7	6.6

3a. What is the mean difference and the SD (not SE) of the difference?

3b. A point halfway between the two means is 3.15 kg. Assuming weight loss follows a normal distribution, what percent of the Zone subjects lost more than 3.15 kg? What percent of the Atkins subjects lost less than 3.15 kg? Does this indicate that there is substantial overlap between the two distributions or did almost all of the Atkins subjects lose more weight than the Zone subjects?

(Note: The EXCEL function =NORMDIST($X, mean, SD, 1$) gives the area behind X for a given mean and SD)

4. The data below compares TB mortality in rats given a placebo vaccine versus rats given a live agent vaccine in a randomized trial

Group	n	num dead	<u>mean</u> follow up in days
Placebo	25	10	7
Agent	25	15	30

4a. Based on statistics that you can derive from this data, which treatment has a lower mortality (hazard) **rate**? (Show work, give units for rate)

4b. In another randomized trial comparing the same two treatments in a new set of rats, survival after $t=14$ days was reported as below. Note that the follow up times were more than 14 days for some or all animals in each group and some animals may die after 14 days. Some died before 14 days. Follow up did NOT end at 14 days. You may assume that the mortality (hazard) **rate** in each group is a constant.

Group	n	pct alive at 14 days (<u>not</u> last follow up time)
Placebo	50	45%
Agent	50	76%

Overall, do the findings of this second study seem to agree or disagree with the findings of the previous study? Show calculations.

5. The hazard rate ratio for post treatment breast cancer (non) survival is $HR=2.0$ in those who drink more than 4 oz of alcohol per day compared to non drinkers. If the 12 month survival in non drinkers is 90%, what is the 12 month survival in alcohol drinkers?

6. Below are serum log 10 IgG (immunoglobulin) titers in children with and without Pertussus.

Titers in those with no Pertussus

0.41, 0.43, 0.47, 0.51, 0.59, 0.67, 0.88, 0.88, 0.91, 0.92, 0.97, 1.06, 1.24, 1.27, 1.27

Titers in those with Pertussus

0.93, 0.95, 1.01, 1.13, 1.17, 1.18, 1.21, 1.22, 1.22, 1.23

We wish to use the serum log IgG level as a medical test for Pertussus. We defined a log IgG level of 1.1 as the threshold. That is, patients with $\log \text{IgG} > 1.1$ test "positive" for Pertussus and those < 1.1 test negative.

Based on the above, compute the estimated sensitivity and specificity of log IgG. You do not have to make any distribution assumptions.

If this data is from a case-control study, should you use this data to compute the positive and negative predictive values? If so, do so. If not, briefly explain why.