Math 2200
Midterm Examination 3 Solutions – April 20, 2010

General Instructions: You may use any calculator you like. You may have a 3x5 card, but no other notes. Only the answer on the answer card will be graded.
Please do not use the binomial model on this exam unless explicitly asked to (even if $np < 10$).

Problems 1-20: Multiple choice. Each problem is worth 4 points.

1. The following table of data gives the calcium concentration in the water of 10 U.K. towns north of Derby, and of ten towns south of Derby.

Derby	Calcium									
North	6	7	8	13	14	15	27	49	71	75
South	5	44	50	59	60	68	90	101	107	133

Is there a difference between the average calcium levels of the two populations? Perform a 2-sided 2-sample t-test and find the P-value.

The problem tells us exactly what test to perform, so we enter the North data into L_1 and the South data into L_2 on our TI, and perform Stat | Tests | 2SampTTest. The reported P-value is D, 0.0082.

2. Perform a (2-sided) Wilcoxon rank-sum test on the data from Question 1 and find the P-value.

We choose the North data to be Sample 1, and add the ranks to get

$$W = 2 + 3 + 4 + 5 + 6 + 7 + 8 + 10 + 15 + 16 = 76.$$

According to the Wilcoxon formulas, $\mu_W = \frac{n_1(n_1+n_2+1)}{2} = 105$ and $\sigma_W = \sqrt{\frac{n_1 n_2(n_1+n_2+1)}{12}} = \sqrt{175} = 5\sqrt{7}$, as $n_1 = n_2 = 10$. This gives a z-score of $\frac{76-105}{13.23} \cong -2.1922$, and our P-value is

$$P(|z| > 2.19) = 2 \cdot \text{normalcdf}(-99, -2.1922) \cong 0.0284,$$

answer D.

3. Perform Tukey's quick test on the data from Question 1 and report the test statistic (the count).

   The smallest value from the pooled sample is 5, and the largest is 144. Since these are both in the South sample, Tukey's quick test does not apply, answer H.

4. A telemarketer is told to expect making a sale on about 10% of calls. What is the probability that, out of 90 calls, she makes exactly 5 sales? Use the binomial model directly.

$$\binom{90}{5} \cdot 0.1^5 \cdot 0.9^{85} \cong 0.0567, \text{ answer D.}$$

5. What is the probability that the telemarketer from Question 4 makes at most 50 sales out of 500 calls? Use the normal approximation for the binomial model, and round to the nearest 0.1.

   The expected number of successful calls is 50, so 50 calls translates to a z-score of 0, and we have a 50% probability of at most this many successes, answer C.

6. Suppose the SAT scores of Washington University in St. Louis students can be described by a normal model with mean 2200 and standard deviation 150 points. What is the probability that the average SAT score of 4 randomly selected Wash U students is at least 2250? (Use the normal model for the sample mean.)

   The normal model for the sample mean has expected value $\mu = 2200$ and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{150}{\sqrt{4}} = 75$. 2250 then translates to a z-score of $\frac{50}{75} = \frac{2}{3}$, and normalcdf$(\frac{2}{3}, 99) \cong 0.2525$, answer E.

7. April showers bring May flowers. Which city receives the most April showers? You estimate that St. Louis on average receives 3.7 inches of rainfall in April, with a margin of error of ± 1.5 inches at the 95% confidence level.
   What is the best description of the meaning of this margin of error?

   Answer E.

8. A Gallup Pool reported that 40% of a random sample of 900 adults said they believe in ghosts. Find the upper bound of a 97% confidence interval for the percentage of adults who believe in ghosts.

   We are finding a confidence interval for a single proportion. On your TI: Stat | Tests | 1PropZInt. Enter $x = 360$, $n = 900$, C-level = .97 – answer C.

9. You are designing a poll to measure the proportion of St. Louis voters who support a certain proposition. Previous polls have shown this proportion hovering near 50%. Rounding up to the next higher thousand, how many voters will you need to sample to ensure a margin of error of 1% (at the 95% confidence level)?

   We solve
   $$.1 = 1.96 \cdot \sqrt{\frac{.5 \cdot .5}{n}}$$
   to get $n = 9604$, which we round up to 10,000, answer C.

10. National data in the 1960s showed that about 45% of the adult population had never smoked cigarettes. A recent poll of 1011 adults found that 52% had never smoked.
    Has there been a change of behavior among Americans? Find the P-value for the appropriate hypothesis test.

    We are testing a single sampled proportion. On your TI: Stat | Tests | 1PropZTest. Enter $p_0 = .45$, $x = 526$, $n = 1011$, and perform a 2-sided test. The resulting P-value is $< .00001$, which rounds to answer H.

11. The Brussel Sprout Research Institute reports that 25% of men report that they enjoy brussel sprouts, while only 20% of women enjoy the tasty cabbage-like vegetable. If these results reflect samples of 240 people of each gender, then is this strong evidence that men and women have different outlooks? What P-value do you find in an appropriate hypothesis test?

    We are comparing 2 sampled proportions. On your TI: Stat | Tests | 2PropZTest, with $x_1 = 60$, $n_1 = 240$, $x_2 = 48$, $n_2 = 240$, 2-sided test. Answer A.

12. In Question 11, the Brussel Sprout Research Institute publishes their estimate that the proportion of men who enjoy brussel sprouts is about 5% higher than the proportion of women. Find the margin of error for this estimate corresponding to the 95% confidence level.

    On your TI: Stat | Tests | 2PropZInt, with parameters as above, C-level of .95. Report half the length of the resulting interval, answer G.

13. In the 1960s, the mean age at which American men first married was 23.3. The following table represents the ages of a random selection of 11 men who married for the first time in 2009.

21	22	22	24	26	26	28	29	32	36	40

    Has the mean age of first marriage changed since 1960? What P-value do you find in an appropriate 2-sided hypothesis test?

    We are testing a single sampled mean. On your TI: Enter data into e.g. L_1. Perform Stat | Tests | T-Test, from data L_1, with $\mu_0 = 23.3$, Freq = 1, 2-sided test. Answer D.

14. In the 1960s, the median age at which American men first married was 22.8. Has the median age of first marriage changed? Using the data from Question 13, what P-value do you find in an appropriate 2-sided sign test?

We count: 3 of 11 are less than 22.8, the remainder greater. On your TI: Stat | Tests | 1PropZTest, $p_0 = .5$, $x = 3$, $n = 11$, 2-sided test. Answer C.

15. To test whether cars get higher mileage per gallon with premium gas (as opposed to regular), we test 8 cars with a tankful of each of regular and premium gasoline. The resulting mileage (in miles per gallon) is below:

Regular	16	19	20	21	23	27	27	28
Premium	19	21	23	23	25	26	29	31

Do cars get better average mileage with premium gasoline? Find the P-value from the appropriate 1-sided hypothesis test.

Since the data represents two measurements on each car, it is a paired t-test. On your TI: Enter Regular row into L_1, and Premium into L_2, then calculate $L_1 - L_2 \to L_3$. Then perform Stat | Tests | T-Test with $\mu_0 = 0$, List L_3, Freq 1, alternate hypothesis $\mu < \mu_0$. Answer G.

16. Consider the data from Question 15 again. Do more than half of car models get better mileage with premium gasoline? Find the P-value from an appropriate 1-sided signed hypothesis test.

We count for a paired sign test: all but 1 of the 8 get better gas mileage. On your TI: Stat | Tests | 1PropZTest, with $p_0 = .5$, $x = 1$, $n = 8$, alternate hypothesis $p < p_0$. (Or symmetrically, $x = 7$, $p > p_0$.) Answer F.

17. After getting trounced by your little brother while playing "Monopoly: the Famous Statisticians Edition", you suspect that the die he gave you was unfair. To check, you roll it 90 times, recording the number of times each face appears:

Face	1	2	3	4	5	6
Count	13	17	11	21	16	12

We are testing a model for the distribution of counts, so use a χ^2 goodness of fit test. On your TI: Enter the counts into L_1, and the expected values of 15, 15, 15, 15, 15, 15 into L_2. The χ^2 statistics is then sum$((L_1 - L_2)^2/L_2) \cong 4.666$, and we look up the P-value with $\chi^2\text{cdf}(4.6666, 10^{99}, 5)$ to find answer D.

18. According to recent polls, about 35% of US voters are Democrats, 32% are Republicans, and 33% are independent. In your own poll of 1000 voters, you find that 330 are Democrats, 328 are Republicans, and 344 are independent. Do these results cast doubt on the 35-32-33 model? Find the P-value of an appropriate χ^2 test.

    We are testing a model for the distribution of counts, so use a χ^2 goodness of fit test.

    Note: it is well-known that mathematicians can't add.

    Assuming that 330+328+344=1000, enter the observed values of 330, 328, and 344 into L_1, and the expected values of 350, 320, and 330 into L_2. As in the previous problem, find sum$((L_1 - L_2)^2/L_2) \cong 1.937$ and $\chi^2\text{cdf}(1.937, 10^{99}, 2)$, answer G.
    Note: if you can add better than a mathematician, then your expected values would be $\frac{350}{1002}$, $\frac{320}{1002}$, and $\frac{330}{1002}$, and the P-value will be off from answer G only slightly.

19. You decide to compare the number of chocolate chips in an 18 oz box of Chip Ahoy! cookies with the number of chocolate chips in an 18 oz box of the supermarket brand. You painstakingly count the number of chips in 6 boxes of each, as summarized in the following table:

Chips Ahoy!	1214	1087	1419	1121	1345	1325
Supermarket	1039	1173	1138	951	1011	1341

    Is $\mu_{\text{Chips Ahoy!}}$ at least 100 greater than $\mu_{\text{Supermarket}}$? Perform an appropriate 1-sided hypothesis test and find the P-value.

    We need to do a 2-sample T test, but we want to compare $\mu_1 - \mu_2$ with 100, rather than 0.

On your TI: Enter Chips Ahoy row into L_1, and Supermarket row into L_2. Reexpress $L_2 + 100 \to L_3$, and perform Stat | Tests | 2SampTTest of L_1 vs L_3 with alternate hypothesis $\mu_1 > \mu_2$. (As usual, Freqs are 1, and Pooled is No.) Answer C.

20. Let X be a random variable with Poisson distribution and $E(X) = 3$. What is $\text{Var}(X)$?

    Answer F.

Problems 21-30: True/false. Each problem is worth 2 points.

21. True/false: After sampling 100 married couples, you want to test the null hypothesis H_0 that men and women earn the same amount. You will use a 2 sample t-test.

    False – the data is paired.

22. True/false: If the probability that a randomly chosen person has tuberculosis is 0.00005, then the number of cases of tuberculosis in greater St. Louis (population about 2.8 million) cannot be modeled with the normal distribution, and must be modeled with the Poisson distribution instead.

    False – $np \cong 140 > 10$.

23. True/false: Purchase amounts at Schnucks grocery stores have a skewed unimodal distribution with a mean of \$35 and standard deviation of \$20. If in a particular day Schnucks sees 312 customers, it is reasonable to model the mean purchase of these 312 customers with a normal distribution.

    True – this is the central limit theorem.

24. True/false: Performing a hypothesis test at a lower significance level will lead to a lower probability of a Type II error.

    False – note that lower significance level = lower α.

25. True/false: Increasing the sample size used in a hypothesis test will lead to a lower probability of a Type I error.

    Either answer was accepted.

26. True/false: A recent study on perfect pitch sampled 2700 students, and found that 7% of non-Asians and 32% of Asians have perfect pitch. A 2-proportion z-test gave a P-value of < 0.0001. We conclude that genetic differences cause the difference in frequency of perfect pitch.

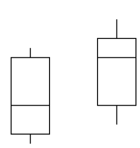
    False – correlation is (still) not causation. (In this case, for example, we may have a lurking variable in that many Asian languages are tonal, and require some precision with pitch.)

27. True/false: A 99% confidence interval for $\hat{p}$ will be wider than the 90% confidence interval for $\hat{p}$.

    True. See the many homework problems on this same topic.

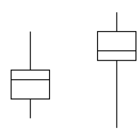
28. True/false: The 99% confidence interval for $\bar{x}$ will be wider than the 90% confidence interval for $\bar{x}$. True, similarly to the previous.

29. True/false: The following boxplot represents two samples of size 40. We conclude from Tukey's quick test that we can reject the null hypothesis (that the means are equal) at a signficance level of 0.001.

True. The Q_3 line of the right sample is above the max of the left sample, and the Q_1 line of the left sample is below the min of the right sample. Since each sample has size 40, we have at least 9 or 10 in both area, for a Tukey statistic of at least 18, which is bigger than 13 (as required).

30. True/false: The following boxplot represents two samples of size 40. We conclude from Tukey's quick test that we can reject the null hypothesis (that the means are equal) at a signficance level of 0.001.



False. Tukey's quick test does not apply, since the right sample has both the overall max and min. (The boxplots are "nested".)