Math 366: Final Exam Review Sheet

Disclaimer: This is not meant to be an exhaustive set of examples. Study all sections we have covered in the book. These are the questions you guys came up with plus a few I added in.

1. A random sample of size $n = 12$ is taken from the distribution with p.d.f. $f(x) = 3x^2,\ 0 \le x \le 1$.
   (a) Find μ and σ^2.
   (b) Find, approximately, $P(\frac{1}{2} \le \overline{x} \le \frac{3}{4})$.
2. Approximate $P(29.03 \le \overline{x} \le 30.79)$, where $\overline{x}$ is the mean of a random sample of size 25 from a distribution with mean $\mu = 30$ and variance $\sigma^2 = 6$.
3. If $X \sim b(80, 0.3)$, find the approximate value of $P(31 \le x \le 34)$ using
   (a) The normal approximation.
   (b) The binomial distribution.
4. A cake factory produces cakes that have a label weight of 54.2 grams. Assume the distribution of the weights of these cakes is $N(55.3, 25)$. Let $\overline{X}$ equal the sample mean of 90 cakes that are selected and weighed on a particular shift. Find $P(55.0 \le \overline{x} \le 55.7)$.
5. Given these 10 data points:

   5 12 4 6 6 5 7 8 1 7

   (a) Create a box and whisker plot and identify any outliers.
   (b) Create a stem and leaf plot.
6. Page 345, Questions 8 and 10.
7. A random sample of $n = 50$ people were asked how often they check their email daily. It was found that $\overline{x} = 5.6, s_X^2 = 8.2$. Give a 90% confidence interval for μ.
8. A survey was done to check the results of the previous question, and the following data was recorded:

   8 6 4 1 3
   7 1 5 2 6

   Find a 90% confidence interval giving an upper bound for μ.
9. An Independent random sample of heights of two kinds of trees yielded the following results, $n = 12$, $\overline{x} = 65.7$ feet, $s_x = 4$ feet, $m = 15$, $\overline{y} = 68.2$ feet, $s_y = 3$ feet. Find an approx 95% CI for the difference $\mu_X - \mu_Y$ of the means of the height, assume $\sigma_X^2 = \sigma_y^2$.
10. Students weighed (in kilos) at the beginning and end of a semester-long health fitness program. Let the random variable D equal the weight change for a student, postweight minus preweight. Assume the distribution of D is $N(\mu_0, \sigma_0^2)$. A random sample of $n = 12$ female students yielded the following observations of D:

    2.0 −0.5 1.4 −2.2 0.3 −0.8
    3.7 −0.1 0.6 0.2 0.9 −0.1

    (a) Give a point estimate of μ_0.
    (b) Find a 95% CI for μ_0.
11. In a class of 150 students, different study methods were used to perpare for an upcoming test: going over a study guide, and reworking problems from sections. Test the null hypothesis that the distributions of scores in each method are equal, given the following data and a significance level of $\alpha = 0.10$.

Grades	F	D	C	B	A	Totals
Methods						
Study Guide	4	17	33	14	11	79
Reworking Problems	0	7	24	29	11	71
Totals	4	24	57	43	22	150

12. Does the number of cups of tea affect the number of hours of sleep at night? Use $\alpha = 0.05$.

Cups	Hours	of	sleep
0	7	8	5
1	6	7	6.5
2	9	8	8
3	9	7	7

13. Does your age have an affect on the number of hours of sleep you get at night? Use $\alpha = 0.05$ and find the p-value (approximate) of this test.

Age	Hours	of	sleep
< 10	9	10	10.5
$10 - 15$	8	9	8.5
$16 - 24$	6	7	7
$25 - 35$	7	7	8

14. Tyler and Trigg report on a mechanic who studied the effect of temperature and time spent welding on the strength of the weld. Below are the results. Three strengths were tested at each combination of temperature and time. Construct an ANOVA table and conduct tests for interaction. Use $\alpha = 0.05$.

	Time	
Temperature	20 min.	40 min.
500°	32, 38, 30	36, 40, 45
550°	40, 37, 35	42, 46, 39
600°	43, 36, 39	39, 47, 48

15. Five different styles of clay pots are tested for their strength when cooked at 3 different heats. The table for their strengths is below:

Heats	A_1	A_2	A_3	A_4	A_5
B_1	44	47	42	51	52
B_2	49	55	43	52	60
B_3	45	48	38	44	56

(a) Use a 5% significance level to test $H_B : \beta_1 = \beta_2 = \beta_3 = 0$ against all alternatives.

(b) Use a 5% significance level to test $H_A : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$ against all alternatives.

16. Let X be the weight of an 8 ounce bag of snow peas from the Waremart produce aisle. Assume $X \sim N(\mu, 0.09)$. We want to test $H_0 : \mu = 8$ against $H_1 : \mu > 8$. Suppose we take a sample of size n and use a critical region of $C = \{\overline{x} : \overline{x} \geq c\}$, and $K(\mu) = P(\overline{x} \geq c; \mu)$. Find c, n such that $\alpha = K(8) = 0.01$ and $K(8.5) = 0.99$.

17. Assume that the number of doughnuts eaten by students during a statistics class has a normal distribution with $\mu = 1.5$ and $\sigma^2 = 1.25$. How many doughnuts should the teacher bring so that the probability of running out of doughnuts is less than 10% in a class of 23 students.

18. Suppose X_1,...,X_n is a random sample of size n from a distribution with pdf $f(x) = \theta x^{\theta-1}$, for $0 \leq x \leq 1$, $0 \leq \theta < \infty$.

(a) Find the Maximum Likelihood Estimator for θ.

(b) Find an estimator for θ using the method of moments.

19. There were 100 independent trials of an experiment run to test $H_0 : \mu = 50$ against $H_1 : \mu > 50$. The critical region was $C = \{\overline{x} : \overline{x} \geq 50.2\}$.

    (a) Write a formula for the power function, $K(\mu)$.

    (b) What is the significance level, α of this test?

    (c) What sample size would be needed if you wanted to make $\alpha = 0.05$? (keeping the same critical region)

    (d) If instead of changing the sample size, we wanted to change the critical region to make $\alpha = 0.05$ (leave $n = 100$) then what would be the new critical region?

20. Cola X and Cola Y are competing in a taste test. People are asked to rate the colas from 1 to 10 with 10 being the best. A survey of 50 people tasting Cola X gave $\overline{x} = 7.6$. A survey of 65 people tasting Cola Y gave $\overline{y} = 8.5$. Assume $\sigma_x^2 = 7.5$ and $\sigma_y^2 = 8.1$. The makers of Cola Y would like to make the claim that they are 95% confident that people like thier cola better. Can they make this claim based on these statistics? Use a confidence interval to justify your answer.

21. A certain medication is supposed to contain 20% of the active ingredient. As a quality control measure, a sample of 12 was taken and the amount of active ingredient measured. The sample mean was $\overline{x} = 20$ and the sample variance was $s^2 = .1$. The company claims the true mean lies in the interval $(19.836, 20.164)$. What level of confidence is there in this interval?

22. In 2000, a poll showed that 79% of children entering Kindergarten had up-to-date immunizations. To update the estimate, how large a sample size is required to be 98% confident that the new estimate has an error no more that $\pm.05$?

23. (a) A 2005 survey showed that 37% of the American viewing audience watched NBC nightly news. If 2000 people were surveyed, find a 90% confidence interval for the proportion of the American viewing audience watching NBC nightly news.

    (b) If NBC wants to repeat the survey to confirm that the percentage still holds, what sample size should they use to maintain the 90% confidence level and have the error be no bigger that $\epsilon = .02$?

24. A sociologist interviews 100 families where the husband and wife, early in their marriage, decide to keep having children until they had their first girl (then stop). (All 100 couples did eventually have a girl.) Assume the probability that any given child is a girl is 0.5. The results were as follows: 55 families had 1 child; 19 families had 2; 12 families had 3; 8 families had 4; 3 families had 5 and 3 families had 6. Do a Chi-Square goodness of fit test to see whether the distribution can be adequately described by a geometric p.d.f. Use $\alpha = 0.05$ and group the families with 5 and 6 children.

Answer the following True or False

25. Having a smaller sample size gives more accurate approximations with using the Central Limit Theorem.

26. When you see "$\leq$" in a binomial distribution question, you should always approximate by subtracting 0.5 from the value; "$<$" would always add 0.5.

27. Using Excel or a calculator is always an acceptable alternative to the $P(n + 1) = I + d$ formula for calculating percentiles.

28. A 90% confidence interval for μ means that 90% of the intervals thus constructed should contain μ.

29. A 90% CI for $\mu_X - \mu_Y$ is $(-10.13, -0.27)$. We can conclude that μ_X is smaller than μ_Y.

30. The $p-$value associated with a test is the probability under H_1 (the alternative hypothesis) that the test statistic is equal to or exceeds the observed value in the direction of H_0 (the null hypothesis).

31. A type I error is defined as the probability of rejecting H_0 and accepting H_1 when the null hypothesis is true.

32. The critical region for $H_0 : \sigma = \sigma_0$ against $H_1 : \sigma < \sigma_0$ is negative.

33. We would reject independence for a Chi-Square experiment with $Q = 7.38, k = 4, \alpha = 0.05$.

34. Using an ANOVA test with $m = 5$ and $n = 35$ and testing $H_0 : \mu_1 = \mu_2 = \cdots = \mu_5$, at an $\alpha = 0.01$ significance level, H_0 is accepted if the computed $F = \dfrac{SS(T)/(5-1)}{SS(E)/(35-5)} \geq 4.02 = F_{0.01}(4, 30)$.

35. $\displaystyle\sum_{i=1}^{m}\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_i)^2 = SS(TO) - SS(T)$.

36. Given $H_0 : \mu = 60$ and $H_1 : \mu > 60$, with critical region $C = \{\overline{x} : \overline{x} \geq 64\}$:

    (a) The probability of making a type I error is $K(64)$.

    (b) $K(57) < \alpha$.

Short Answer

37. The Central Limit Theorem is used to approximate a distribution as if it were $\sim N(\_, \_)$.

38. The Normal is a good approximation to the Binomial if $np \geq$ _____?

39. What is the difference between and outlier and a suspected outlier?

40. As the sample size, n, increases and α remains the same, the confidence interval gets (larger / smaller).

41. A 95% CI for $\mu_X - \mu_Y$ is $(-3.23, 0.02)$. We can conclude that _____.

42. What table should you use to test the mean if σ^2 is unknown?

43. Suppose you perform a Chi-Square goodness of fit test and find your test statistic to be $q_5 = 12.01$; using $\alpha = 0.05$, we will _____ the null hypothesis. (accept/reject)

44. What level of significance is needed to confirm that a given distribution is $b(n, p)$ using a Chi-Square goodness of fit test if the experiment has 4 outcomes and the test statistic is $Q = 8.13$? That is, what is the highest value of α that will accept H_0?

45. The test statistic using a contingency table for an experiment testing the heart rate of adults grouped by age $(18-30, 31-55, 55+)$ in relation to the corner of a room in which they are placed (blue, red, black, yellow, green) has _____ degrees of freedom.

Answers

1. (a) $\mu = \frac{3}{4}$; $\sigma^2 = \frac{3}{80}$.
   (b) 0.5
2. 0.9224
3. (a) 0.0507
   (b) 0.0524
4. 0.4921
5. (a) The 5-number summary is $(1, 4.75, 6, 7.25, 12)$. No outliers.
   (b) See your book if you don't know how to do this.
6. Problem 10: a) $\hat{p} = \frac{1}{\overline{x}}$; b) $\hat{p}$=number of success/number of trial c) 0.0794
7. $(4.93, 6.27)$
8. $(0, 5.28)$
9. $(-5.273, 0.273)$
10. (a) 0.45
    (b) $(-0.32, 1.22)$
11. $q_4 = 14.43$; $\chi^2_{0.10}(4) = 7.79$; Reject H_0
12. $F = 2.11$, $F_{0.05}(3, 8) = 4.09$, do not reject H_0
13. $F = 9.13$, $F_{0.05}(3, 8) = 4.07$, Reject H_0
14. $F_{AB} = 0.111$, (no interaction); $F_A = 2.6$ (no row effect); $F_B = 9.69$ (Yes, column effect)
15. (a) $F_B = 15.5$ - reject H_B
    (b) $F_A = 7.62$ - reject H_A
16. $c = 8.25$, $n = 8$
17. 42
18. (a) $\frac{-n}{ln(\prod_{i=1}^{n} X_i)}$
    (b) $\frac{\overline{X}}{1-\overline{X}}$
19. (a) $K(\mu) = 1 - \Phi(\frac{50.2-\mu}{(1/5)})$
    (b) $\alpha = 0.1587$
    (c) $n = 271$
    (d) $C = \{\overline{x} : \overline{x} \geq 50.329\}$
20. The confidence interval for $\mu_X - \mu_Y$ is $(-1.93, 0.127)$ and in particular contains 0. Hence the Cola Y makers cannot claim they are 95% confident that their cola is better since their data shows it could be better, worse or the same.
21. 90% (Note: use $t-$table.)
22. 360
23. (a) $(0.3522, 0.3878)$

(b) 1577

24. $q_4 = 2.46$ accept H_0 (geometric p.d.f. describes the distribution)

TRUE/FALSE

25. FALSE
26. FALSE
27. FALSE
28. TRUE
29. TRUE
30. FALSE
31. TRUE
32. FALSE
33. FALSE
34. FALSE
35. TRUE
36. (a) FALSE
    (b) TRUE

Short answer

37. $N(0, 1)$
38. 5
39. see book p. 330
40. smaller
41. Nothing (since 0 is in the CI)
42. t-table
43. reject
44. $\alpha = 0.05$
45. 8