

## AP Exam Ch. 7-12 Review

### Chapter 7: Sampling Distributions

1. A certain population of male whooping cranes that migrate between Wisconsin and Florida every year has a SRS taken. The mean weight of all males is assumed to be 22.7 pounds with a standard deviation of 2.3 pounds and is said to follow a Normal distribution.
  - a. What is the probability that a randomly selected male crane weighs less than 20 pounds? Sketch the curve and put in all the appropriate values.
  - b. When these cranes reach Florida, a random sample of 25 male cranes is weighed and measured. What is the probability that the mean weight of the sample will be between 15 and 22 pounds?
  - c. What is the probability that the mean weight of a sample of 25 cranes would be greater than 23 pounds?
  - d. What is the probability that the mean weight of 25 cranes would be less than 18 pounds?
2. The Helsinki Heart Study asks whether the anti-cholesterol drug gemfibrozil will reduce heart attacks. In planning such an experiment, the researchers must be confident that the sample sizes are large enough to enable them to observe enough heart attacks. The Helsinki study plans to give gemfibrozil to 2000 men and a placebo to another 2000 men. The probability of a heart attack during the 5-year period of the study for men this age is about 0.04. We can think of the study participants as an SRS from a large population, of which the proportion  $p = 0.04$  will have heart attacks.
  - a. What is the mean number of heart attacks that the study will find in one group of 2000 men if the treatment doesn't change the probability of 0.04?
  - b. What is the probability that the group will suffer at least 75 heart attacks? Sketch the curve, show all the work and write the probability statement.

3. Children in kindergarten are sometimes given the Raven Progressive Matrices Test (RPMT) to assess their readiness for learning. Experience at Southward Elementary School suggests that the RPMT scores for its kindergarten pupils have a mean of 13.6 with a standard deviation of 3.1. The distribution is close to Normal. Mr. Brown has 22 children in his kindergarten class this year.
- What is the probability that class's mean score will be less than 12.0?
  - Mr. Brown suspects that the class RPMT scores will be unusually low because the test was interrupted by a fire drill. He wants to find the level  $L$  such that there is only a probability of 0.05 that the mean score of his class fall below  $L$ . What is this value of  $L$ . (Hint: this requires you to find the  $z$ -score and then convert to the  $x$ -score.)
4. What is the Central Limit Theorem? How is the CLT used in sampling distributions?

### Chapter 8-9: Estimating with Confidence and Testing a Claim

#### 5. Estimating Population Parameters.

- Why is an unbiased statistic generally preferred over a biased statistic for estimating a population characteristic?
- Does unbiasedness alone guarantee that the estimate will be close to the true value? Explain.
- A random sample of 12 four-year old red pine trees was selected and the diameter (in) of each tree's main stem was measured. The resulting observations are as follows:

11.3   10.7   12.4   15.2   10.1   12.1   16.2   10.5   11.4   11.0   10.7   12.0

Find the estimate that can be used to estimate the true population mean.

Find the estimate that can be used to estimate the true population standard deviation.

Find the estimate that be used to estimate the true population proportion of trees whose diameter is greater than the average.

6. What is meant by the standard error of a population parameter? What are the standard errors for the following:

Population Mean

Population Proportion

Difference between Two Population Means

Difference between Two Population Proportions

7. What is the general form of all confidence intervals?
8. Suppose that a random sample of 50 bottles of a particular brand of cough medicine is selected and the alcohol content of each bottle is determined. Let  $\mu$  denote the average alcohol content for the population of all bottles of the brand under the study. Suppose that the sample mean is 8.2 grams with a sample standard deviation of 1.5 grams.
- Construct and interpret a 95% confidence interval for the mean alcohol content of the cough medicine.
  - Would a 90% confidence interval be narrower or wider? Explain.
  - The manufacturer claims that the alcohol content is 8.0 grams per bottle. Perform a hypothesis test to test the manufacturer's claim.

9. Retailers report that the use of cents-off coupons is increasing. The Scripps Howard News Service reported that proportion of all households that use coupons as 0.77. Suppose that this estimate was based on a random sample of 800 households.

a. Construct and interpret a 99% confidence interval for the true population proportion.

b. The manager of the retail store in reporting to his superiors claims that the true proportion of customers that use coupons is 80%. Test the manager's claim.

10. Explain the relationship between the t-test and the z-test in hypothesis testing.

11. What is meant by margin of error in a confidence interval?

### Chapter 10: Comparing Two Populations or Groups

12. Are girls less inclined to enroll in science courses than boys? One recent study asked randomly selected 4<sup>th</sup>, 5<sup>th</sup> and 6<sup>th</sup> graders how many science courses they intend to take in high school. The following data was obtained:

	n	Mean	Standard Deviation
Males	203	3.42	1.49
Females	224	2.42	1.35

a. Construct and interpret a 99% confidence interval for the difference between males and females in mean number of science courses planned.

b. The science teacher at this high school claims that there is no difference in the number of courses boys and girls take. Test the science teacher's claim.

**13.** Techniques for processing poultry were examined by a manufacturer of canned chicken. Whole chickens were chilled 0, 2, 8 and 24 hours before being cooked and canned. To determine whether the chilling time affected the texture of the canned chicken, samples were evaluated by trained testers. One characteristic of interest was hardness. Each mean is based on 36 ratings.

Chilling Time

	0 hour	2 hour	8 hour	24 hour
Mean Hardness	7.52	6.55	5.70	5.65
Standard Deviation	.96	1.74	1.32	1.50

a. Does the data suggest that there is a difference in mean hardness for chicken chilled 0 hours before cooking and chicken chilled 2 hours before cooking? Just do the calculations.

b. Does the data suggest that there is a difference in mean hardness for chicken chilled 8 hours before cooking and chicken chilled 24 hours before cooking? Just do the calculations.

c. If a Type I error were made in part a, what would this mean? What are the consequences?

d. If a Type II error were made in part b, what would this mean? What are the consequences?

**14.** The discharge of industrial wastewater into rivers affects water quality. To assess the effect of a particular power plant on water quality, 24 water specimens were taken 16 km upstream and 4 km downstream of the plant. Alkalinity (mg/L) was determined for each specimen, resulting in the summary quantities in the table below.

Location	n	Mean	Standard Deviation
Upstream	24	75.9	1.83
Downstream	24	183.6	1.70

a. Does the data suggest that the true mean alkalinity is higher downstream than upstream by more than 50 mg/L? Just do the calculations.

b. Find the 90% confidence interval for the mean difference. Just do the calculations. Does this confirm your conclusion in the hypothesis test?

15. The article “Softball Sliding Injuries” provided a comparison of breakaway bases (designed to reduce injuries) and stationary bases. Consider the accompanying data. Does the use of breakaway bases reduce the proportion of games in which a player suffers a sliding injury? Perform a test at the 1% significance level.

	Number of Games Played	Number of Games Where a Player Suffered a Sliding Injury
Stationary Bases	1250	90
Breakaway Bases	1250	20

### Chapter 11: Inference for Distributions of Categorical Data

16. The color vision of birds plays a role in their foraging behavior. Birds use color to select and avoid certain types of food. The authors of the article “Color Avoidance in Northern Bobwhites” studied the pecking behavior of 1-day-old bobwhites. In an area painted white, they inserted four pins with different colored heads. The color of the pin chosen on the birds’ first peck for each of 33 bobwhites, resulting in the accompanying table. Does this data provide evidence of color preference?

Color	Blue	Green	Yellow	Red
First Peck Frequency	16	8	6	3

17. Do women have different patterns of work behavior than men? The article “Workaholism in Organizations: Gender Differences” attempts to answer this question. Each person in a random sample of 423 graduates of a business school in Canada were polled and classified by gender and workaholism type, resulting in the accompanying table:

- a. Test the hypothesis that gender and workaholism type are independent.

Workaholism	Female	Male
Work Enthusiasts	20	41
Workaholics	32	37
Enthusiastic Workaholics	34	46
Unengaged Workers	43	52
Relaxed Workers	24	27
Disenchanted Workers	37	30

- b. The author writes “women and men fell into each of the six workaholism types to a similar degree.” Does the outcome of the test you performed in part a support this conclusion? Explain.

## Chapter 12: More About Regression

18. It is certainly plausible that workers are less likely to quit their jobs when wages are high than when they are low. The paper “Investigating the Causal Relationship Between Quits and Wages” presented the accompanying data on  $x$  = average hourly wages and  $y$  = quit rate (number of employees per 100 who left jobs during 1986.) Each observation is for a different industry.

x	8.20	10.35	6.18	5.37	9.94	9.11	10.59	13.29	7.99	5.54	7.50	6.43	8.83	10.93	8.80
y	1.4	.7	2.6	3.4	1.7	1.7	1.0	.5	2.0	3.8	2.3	1.9	1.4	1.8	2.0

The following is the Minitab output:

Predictor	Coef	Stdev	t-ratio	p
Constant	4.8615	0.5201	9.35	0.0000
Wage	-0.34655	0.05866	-5.91	0.0000

$s = 0.4862$                        $R\text{-sq} = 72.9\%$                        $R\text{-sq(adj)} = 70.8\%$

Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	1	8.2507	8.2507	34.90	0.0000
Error	13	3.0733	0.2364		
Total	14	11.3240			

- Identify and interpret the slope and y-intercept for the LSRL for average hourly wages and quit rate.
- Write the equation of the LSRL. Define any variables used in the equation.
- Construct and interpret the 95% confidence interval for the slope of the line.

- d. Test the hypothesis that there is a negative relationship between average hourly wages and quit rate.