

1. (5 points) **Substance M** In an experiment to learn if substance M can help restore memory, the brains of 20 rats were treated to damage their memories. The rats were trained to run a maze. After a day, 10 rats were given M and 7 of them succeeded in the maze; only 2 of the 10 control rats were successful. The z test for “no difference” against “a higher proportion of the M group succeeds” has

A. $z = 2.60, p < 0.005$.

B. $z = 2.25, 0.02 < p < 0.04$.

C. $z = 2.25, p < 0.02$.

2. (5 points) **Alaskan Smokers** In the past decade there have been intensive antismoking campaigns sponsored by both federal and private agencies. The Behavioral Risk Factor Surveillance System (BRFSS) is an ongoing data collection program that monitors behaviors such as smoking on a statewide level using data collected on a random sample of adults through a telephone survey. The first sample, taken in 1999 in Alaska, involved 2045 adults, of which 592 were current smokers. The second sample, taken in 2009 in Alaska, involved 2411 adults, of which 511 were smokers. The samples are to be compared to determine whether the proportion of U.S. adults in Alaska that smoke declined during the 10-year period between the samples. Take p_{1999} and p_{2009} to be the proportions of all adults in Alaska over 18 who were current smokers in these years. The sample proportions of adults who were current smokers in 1999 and 2009 are about

A. $\hat{p}_{1999} = 0.21$ and $\hat{p}_{2009} = 0.25$.

B. $\hat{p}_{1999} = 0.21$ and $\hat{p}_{2009} = 0.29$.

C. $\hat{p}_{1999} = 0.29$ and $\hat{p}_{2009} = 0.21$.

3. (5 points) Because the t procedures are robust, the most important condition for their safe use is that

A. the data can be regarded as an SRS from the population.

B. the population standard deviation σ is known.

C. the population distribution is exactly Normal.

4. (5 points) One major reason that the two-sample t procedures are widely used is that they are quite robust. This means that

A. the t procedures do not require that we know the standard deviations of the populations.

B. the t procedures compare population means, a comparison that answers many practical questions.

C. confidence levels and P-values from the t procedures are quite accurate even if the population distribution is not exactly Normal.

5. (20 points) Hallux abducto valgus (call it HAV) is a deformation of the big toe that often requires surgery. Doctors used X-rays to measure the angle (in degrees) of deformity in 38 consecutive patients under the age of 21 who came to a medical center for surgery to correct HAV. The angle is a measure of the seriousness of the deformity. Here are the data:

28	32	25	34	38	26	25	18	30	26	28	13	20
21	17	16	21	23	14	32	25	21	22	20	18	26
16	30	30	20	50	25	26	28	31	38	32	21	

A good way to judge the effect of an outlier is to do your analysis twice, once with the outlier and a second time without it. The data above follow a Normal distribution quite closely except for one patient with HAV angle 50 degrees, a high outlier.

(a) It is reasonable to regard these patients as a random sample of young patients who require HAV surgery?

(b) Find the 95% confidence interval for the population mean based on the 38 patients who remain BEFORE you drop the outlier.

(c) Find the 95% confidence interval for the population mean based on the 37 patients who remain AFTER you drop the outlier. Compare your interval in (b) with this interval. What is the most important effect of removing the outlier?

6. (20 points) “Durable press” cotton fabrics are treated to improve their recovery from wrinkles after washing. Unfortunately, the treatment also reduces the strength of the fabric. A study compared the breaking strength of untreated fabric with that of fabrics treated by three commercial durable press processes. Five specimens of the same fabric were assigned at random to each group. Here are the data, in pounds of pull needed to tear the fabric:

Untreated	60.1	56.7	61.5	55.1	59.4
Permafresh 55	29.9	30.7	30.0	29.5	27.6
Permafresh 48	24.8	24.6	27.3	28.1	30.3
Hylite HF	28.8	23.9	27.0	22.1	24.2

The untreated fabric is clearly much stronger than any of the treated fabrics. We want to know if there is a significant difference in breaking strength among the three durable press treatments.

(a) Analyze the data for the THREE PROCESSES by carrying out an ANOVA test.

(b) Which process do you recommend if breaking strength is a main concern? (Although the standard deviations do not quite satisfy our rule of thumb, that rule is conservative, and many statisticians would use ANOVA for these data.)

7. (20 points) Subjects with preexisting cardiovascular symptoms who were receiving subitramine, an appetite suppressant, were found to be at increased risk of cardiovascular events while taking the drug. The study included 9804 overweight or obese subjects with preexisting cardiovascular disease and/or type 2 diabetes. The subjects were randomly assigned to subitramine (4906 subjects) or a placebo (4898 subjects) in a double-blind fashion. The primary outcome measured was the occurrence of any of the following events: nonfatal myocardial infarction or stroke, resuscitation after cardiac arrest, or cardiovascular death. The primary outcome was observed in 561 subjects in the subitramine group and 490 subjects in the placebo group. Do the data give good reason to think that there is a difference between the proportions of treatment and placebo subjects who experienced the primary outcome? (Note that subitramine is no longer available in the United States due to the manufacturer's concerns over increased risk of heart attack or stroke.)

State hypotheses, find the test statistic, and use either software or the bottom row of Table C for the P-value. Be sure to state your conclusion.

8. (20 points) Kathleen Vohs of the University of Minnesota and her coworkers carried out several randomized comparative experiments on the effects of thinking about money. Here's part of one such experiment. Ask student subjects to unscramble 30 sets of five words to make a meaningful phrase from four of the five words. The control group unscrambled phrases like "cold it desk outside is" into "it is cold outside." The treatment group unscrambled phrases that lead to thinking about money, turning "high a salary desk paying" into "a high-paying salary." Then each subject worked a hard puzzle, knowing that he or she could ask for help. Here are the times in seconds until subjects asked for help.

For the treatment group:

609	444	242	199	174	55	251	466	443
531	135	241	476	482	362	69	160	

For the control group:

118	272	413	291	140	104	55	189	126
400	92	64	88	142	141	373	156	

The researchers suspected that money is connected with self-sufficiency, so that the treatment group will ask for help less quickly (will go longer without asking for help) on the average. Do the data support this idea?