One-tailed Z-test

8.1-8 It was claimed that 75% of all dentists recommend a certain brand of gum for their gum-chewing patients. A consumer group doubted this claim and decided to test H_0: $p = 0.75$ against the alternative hypothesis H_1: $p < 0.75$, where p is the proportion of dentists who recommend this brand of gum. A survey of 390 dentists found that 273 recommended this brand of gum. Which hypothesis would you accept if the significance level is

(a) $\alpha = 0.05$?
(b) $\alpha = 0.01$?
(c) Find the p-value for this test.

8.1-9 It was claimed that the proportion of Americans who select jogging as one of their recreational activities is $p = 0.25$. A shoe manufacturer thought that p was larger than 0.25. They decided to test the null hypothesis H_0: $p = 0.25$ against the alternative hypothesis H_1: $p > 0.25$. If $y = 1497$ out of a random sample of $n = 5757$ selected jogging, what is your conclusion at a significance level of

(a) $\alpha = 0.05$?
(b) $\alpha = 0.025$?
(c) Find the p-value for this test.

Two-proportions Z-test (Section 14 in notes)

8.1-17 A machine shop that manufactures toggle levers has both a day and a night shift. A toggle lever is defective if a standard nut cannot be screwed onto the threads. Let p_1 and p_2 be the proportion of defective levers among those manufactured by the day and night shifts, respectively. We shall test the null hypothesis, H_0: $p_1 = p_2$, against a two-sided alternative hypothesis based on two random samples, each of 1000 levers taken from the production of the respective shifts.

(a) Define the test statistic and a critical region that has an $\alpha = 0.05$ significance level. Sketch a standard normal p.d.f. illustrating this critical region.
(b) If $y_1 = 37$ and $y_2 = 53$ defectives were observed for the day and night shifts, respectively, calculate the value of the test statistic. Locate the calculated test statistic on your figure in part (a) and state your conclusion.

8.1-18 *Time, April 18, 1994*, reported the results of a telephone poll of 800 adult Americans, 605 of them nonsmokers, who were asked the following question: "Should the federal tax on cigarettes be raised by \$1.25 to pay for health care reform?" Let p_1 and p_2 equal the proportions of nonsmokers and smokers, respectively, who would say yes to this question. Given that $y_1 = 351$ nonsmokers and $y_2 = 41$ smokers said yes,

(a) With $\alpha = 0.05$, test H_0: $p_1 = p_2$ against H_1: $p_1 \neq p_2$.
(b) Find a 95% confidence interval for $p_1 - p_2$. Is this in agreement with the conclusion of part (a)?
(c) Find a 95% confidence interval for p, the proportion of adult Americans who would say yes.

8.1-19 Let p_m and p_f be the respective proportions of male and female white-crowned sparrows that return to their hatching site. Give the endpoints for a 95% confidence interval for $p_m - p_f$, given that 124 out of 894 males and 70 out of 700 females returned. (*The Condor*, 1992, pp. 117–133.) Does this agree with the conclusion of a test of H_0: $p_1 = p_2$ against H_1: $p_1 \neq p_2$ with $\alpha = 0.05$?

Good-of-fit Q-test (Section 17 in notes)

8.5-4 In a large bin of crocus bulbs it is claimed that 1/4 will produce yellow crocuses, 1/4 will produce white crocuses, and 1/2 will produce purple crocuses. If 40 bulbs produced 6 yellow, 7 white, and 27 purple crocuses, would the claim be rejected at the $\alpha = 0.05$ significance level?

8.5-5 In a biology laboratory the mating of two red-eyed fruit flies yielded $n = 432$ offspring, among which 254 were red-eyed, 69 were brown-eyed, 87 were scarlet-eyed, and 22 were white-eyed. Use these data to test, with $\alpha = 0.05$, the hypothesis that the ratio among the offspring would be 9:3:3:1, respectively.

8.5-11 While testing a used tape for bad records, a computer operator counted the number of flaws per 100 feet of tape. Let X equal this random variable. Test the null hypothesis that X has a Poisson distribution with a mean of $\lambda = 2.4$ given that 40 observations of X yielded 5 zeros, 7 ones, 12 two, 9 threes, 5 fours, 1 five, and 1 six. Let $\alpha = 0.05$.
HINT: Combine five and six into one set; that is, the last set would be all x values ≥ 5.

Contingency tables (Section 18 in notes)

8.6-7 A random sample of 100 students were classified by gender and by the "instrument" that they "played." Test whether the selection of instrument is independent of the gender of the respondent. Approximate the p-value of this test.

	Instrument					
Gender	Piano	Woodwind	Brass	String	Vocal	Totals
Male	4	11	15	6	9	45
Female	7	18	6	6	18	55
Totals	11	29	21	12	27	100

8.6-8 A student who uses the college recreational facilities was interested in whether there is a difference between the facilities used by men and women. Test the null hypothesis that facility and gender are independent attributes using the following data. Use $\alpha = 0.05$.

	Facility		
Gender	Racquetball Court	Track	Totals
Male	51	30	81
Female	43	48	91
Totals	94	78	172

Anova (one-factor) F-test (Section 19)

8.7-3 Four groups of three pigs each were fed individually four different feeds for a specified length of time to test the hypothesis H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$, where μ_i is the mean weight gain for each of the feeds, $i = 1, 2, 3, 4$. Determine whether the null hypothesis is accepted or rejected at a 5% significance level if the observed weight gains are, respectively,

X_1:	194.11	182.80	187.43
X_2:	216.06	203.50	216.88
X_3:	178.10	189.20	181.33
X_4:	197.11	202.68	209.18

8.7-4 For the following set of data show that the computed $SS(E)/(n-m) = 1$ and $SS(T)/(m-1) = 75$. This suggests that the unbiased estimate of σ^2 based on SS(T) is usually greater than σ^2 when the true means are unequal.

X_1:	4	5	6
X_2:	9	10	11
X_3:	14	15	16

Anova (two-factor) F-test (Section 20)

8.8-3 We wish to compare compressive strengths of concrete corresponding to $a = 3$ different drying methods (treatments). Concrete is mixed in batches that are just large enough to produce three cylinders. Although care is taken to achieve uniformity, we expect some variability among the $b = 5$ batches used to obtain the following compressive strengths. (There is little reason to suspect interaction and hence only one observation is taken in each cell.)

	Batch				
Treatment	B_1	B_2	B_3	B_4	B_5
A_1	52	47	44	51	42
A_2	60	55	49	52	43
A_3	56	48	45	44	38

(a) Use the 5% significance level and test H_A: $\alpha_1 = \alpha_2 = \alpha_3 = 0$ against all alternatives.

(b) Use the 5% significance level and test H_B: $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$ against all alternatives. [See R. V. Hogg and J. Ledolter, *Applied Statistics for Engineers and Physical Scientists*, 2nd ed., (New York: Macmillan, 1992).]

8.8-4 Show that the cross-product terms formed from $(X_{ij} - \overline{X}_{i\cdot} - \overline{X}_{\cdot j} + \overline{X}_{\cdot\cdot})$, $(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})$, and $(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})$, sum to zero, $i = 1, 2, \ldots a$ and $j = 1, 2, \ldots, b$.
HINT: For example, write

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})(X_{ij} - \overline{X}_{i\cdot} - \overline{X}_{\cdot j} + \overline{X}_{\cdot\cdot})$$

$$= \sum_{j=1}^{b}(\overline{X}_{\cdot j} - \overline{X}_{\cdot\cdot})\sum_{i=1}^{a}[(X_{ij} - \overline{X}_{\cdot j}) - (\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})]$$

Additional problems:

① Show that Z (from p.56 in notes), when squared, equals Q (on p. 78) in the case $k=2$.
Hint: use the arguments on p. 69 in the notes and apply this to both <u>rows</u> AND <u>columns</u>. Also use

$$n_1\left(\frac{x_1}{n_1} - \hat{p}\right) = -n_2\left(\frac{x_2}{n_2} - \hat{p}\right)$$

② Calculate the expectation of an $F(d_1, d_2)$ distribution.
Use that the density of a χ^2(k d.o.f.) has density

$$f_{\chi^2(k)}(x) = x^{k/2-1}e^{-x/2} \Big/ 2^{k/2}\,\Gamma(k/2)$$