Corollary 4.4 (to Theorem 4.15), p. 155, states that for independent \( X_1, \ldots, X_n \),
\[
\text{cov} \left( \sum_{i=1}^{n} a_i X_i, \sum_{i=1}^{n} b_i X_i \right) = \sum_{i=1}^{n} a_i b_i \text{var} (X_i).
\]
Apply this result with \( a_i = -1/n \) for \( i \neq r \), \( a_r = 1 - 1/n \), and \( b_i \equiv 1/n \).

Thus the deviations of the observations from the sample mean are uncorrelated with the sample mean. No conditions on the distribution are assumed, other than existence of finite second moments.

If, further, the distribution is symmetric about its mean, then it also is true that the sample variance is uncorrelated with the sample mean.

If, still further, the distribution is normal, then the sample variance and the sample mean are independent, as we see in Theorem 8.11.

We discuss this in class.

This concerns a random sample of size \( n \) (without replacement) from the discrete uniform distribution on \( 1, 2, \ldots, N \). For this distribution, the mean \( \mu \) and variance \( \sigma^2 \) are given in Exercise 5.1. Use these in Theorem 8.6 to obtain the given answers for the mean and variance of \( \bar{X} \).

We discuss this in class.

Write \( \sum_{i=1}^{n}(X_i - \mu)^2 = \sum_{i=1}^{n}[(X_i - \bar{X}) + (\bar{X} - \mu)]^2 \) and proceed in the obvious way.

Figure out the mean and variance of the chi-square distribution with 1 degree of freedom, and then apply the central limit theorem, Theorem 8.3.

Apply the fact that a sum of independent chi-square random variables is another chi-square random variable, with degrees of freedom equal to the sum of those for the summands.

Let \( Y \) be chi-square with \( \nu = 50 \). Then (justify steps!)
\[
P(Y > 68.0) = P \left( \frac{Y - 50}{\sqrt{100}} > \frac{68 - 50}{10} \right) \approx P(N(0, 1) > 1.8) = 0.036.
\]

Nice connection, isn’t it?

Write
\[
1 / \left( \frac{U/\nu_1}{V/\nu_2} \right) = \frac{V/\nu_2}{U/\nu_1}
\]
and apply definitions.

Apply Theorem 8.16 with \( r = m, n = 2m + 1 \), and \( f \) the density function of uniform(0,1).
8.77 Let’s translate this into mathematical statistics. A claim that a normal population has variance \( \sigma^2 = 4 \) is to be rejected if the sample variance \( s^2 \) based on a sample of size \( n = 9 \) from this population exceeds a threshold of 7.7535. We want to find the probability that the claim becomes rejected when in fact it is true.

If the claim is true, then by Theorem 8.11 the random variable

\[
\frac{(n - 1)s^2}{\sigma^2} = \frac{(9 - 1)s^2}{4} = 2s^2
\]

has a chi-square distribution with \( n - 1 = 8 \) degrees of freedom. Now use Table V, p. 576, as illustrated in Figure 8.1, p. 277, and find

\[
P(s^2 > 7.7535) = P(2s^2 > 15.507) = P(\chi^2_8 > 15.507) = ??
\]

We would hope that this probability is small – right? The answer in the text, 0.50, is incorrect, by the way. What answer do you find?

8.78 Set this up in similar fashion as in Problem 8.78 and use Table IV with the \( t \) distribution with 24 degrees of freedom.

9.7 We discuss this in class.

9.28 This is an estimation problem, so the set of actions is the same as the set of states of nature: \{0, 1/2, 1\}. Since this set has 3 elements, there are \( 3 \times 3 = 9 \) pairs of arguments \((a, \theta)\) for the loss function.

(a) Show the loss function.

(b) Since the data \( X \) is two-valued (0 or 1), and for each value there are 3 choices of action, there are \( 3^2 = 9 \) different decision functions that can be defined. List these decision functions.

(Incidentally, the appearance of ”9” in both (a) and (b) is a coincidence.)

(c) Now make a table showing, for each decision rule, the (three-valued) risk function defined over \( \theta = 0, 1/2, 1 \). Then identify and eliminate five decision rules which are inadmissible. For the other four decision rules, evaluate the maximum risk for each and see that they all are minimax (with maximum risk of 50, each).

(d) Now assume a probability distribution on \( \theta \) placing probability 1/3 on each of the values 0, 1/2, and 1. Calculate the Bayes risks and find the one decision rule with minimum Bayes risk = 16 2/3. This is the rule that chooses action \( a = 0 \) if \( X = 0 \) and \( a = 1 \) if \( X = 1 \). It also is one of the four minimax rules.

10.14 Using the hint, consider \( X \) as the sum of \( n \) Bernoulli random variables \( X_i \) each having density \( f(x; \theta) = \theta^x(1 - \theta)^{1-x} \) for \( x = 0 \) or 1. Then \( X/n = \bar{X} \) is unbiased for \( \theta \). Now check that

\[
\ln f(x; \theta) = x \ln \theta + (1 - x) \ln(1 - \theta)
\]

and

\[
\frac{\partial \ln f(x; \theta)}{\partial \theta} = \frac{x - \theta}{\theta(1 - \theta)}.
\]

Then evaluate the Cramér-Rao lower bound and find that it equals \( \theta(1 - \theta)/n \), which is attained by the variance of \( X/n \).
10.15 Consider $X_1, \ldots, X_n$ to be a random sample from Poisson($\lambda$). The relevant density is thus $f(x; \lambda) = \lambda^x e^{-\lambda}/x!$ for $x = 0, 1, \ldots$, and $\overline{X}$ is unbiased and has variance $\lambda/n$. Now check that 
\[
\frac{\partial \ln f(x; \lambda)}{\partial \lambda} = \frac{x}{\lambda - 1}
\]
and obtain the Cramér-Rao lower bound.

10.17 Proceed in the same way as with 10.14 and 10.15 :-)

10.33 Suppose here a sample of size $n$. For any $c > 0$ (and without loss of generality < 1),
\[
P(|Y_1 - \alpha| < c) = P(\alpha - c < Y_1 < \alpha + c) = P(\alpha < Y_1 < \alpha + c) = 1 - P(\text{all } n \text{ observations are greater than } \alpha + c) = 1 - (1 - c)^n \to 1, \ n \to \infty.
\]
(Be able to justify the above steps.)

10.42 Here assume that $X$ has the exponential distribution with parameter $\theta$, given by the density
\[
f(x; \theta) = (1/\theta)e^{-x/\theta}, x > 0.
\]
(Recall Definition 6.3.) The joint density is
\[
f(x_1, \ldots, x_n; \theta) = (1/\theta)^n e^{-n\sum x_i/\theta},
\]
and it is straightforward to apply Theorem 10.4, the factorization theorem.

10.48 Show using Definition 5.5 that the joint mass function $f(x_1, \ldots, x_n)$ of $n$ independent geometric($\theta$) random variables can be written as
\[
\theta^n (1 - \theta)(\sum_{i=1}^n x_i)^{-n}
\]
and apply Theorem 10.4, the factorization theorem.

10.57 Since there are two parameters $\alpha$ and $\beta$ in uniform($\alpha, \beta$) to be estimated, express the first two moments $\mu'_1 = E(X)$ and $\mu'_2 = E(X^2)$ in terms of $\alpha$ and $\beta$ and solve for $\alpha$ and $\beta$. Then, in the solution equations, substitute the sample moments $m'_1$ and $m'_2$.
Just for fun, compare with the result of Example 10.13 treating the method-of-moments estimator of $\alpha$ in uniform($\alpha, 1$).

10.63 Enjoy!

10.77 We have
\[
f(x|\Lambda) = \Lambda^x e^{-\Lambda}/x!, \ x = 0, 1, \ldots,
\]
and
\[
g(\Lambda) = c \Lambda^{\alpha-1} e^{-\Lambda/\beta}, \ \Lambda > 0.
\]
Then the joint distribution of $X$ and $\Lambda$ is of form

$$c \frac{1}{x!} \Lambda^{x+\alpha-1} e^{-\Lambda(1+1/\beta)},$$

and hence after dividing by the marginal density of $X$ (which can depend on $\alpha$ and $\beta$ but not on $\Lambda$) we obtain the conditional distribution of $\Lambda$ given $X = x$:

$$h(x, \alpha, \beta) \Lambda^{x+\alpha-1} e^{-\Lambda(1+1/\beta)},$$

for some function $h(x, \alpha, \beta)$ not depending on $\Lambda$. Now obtain (a) and (b) easily.

10.87 Straightforward calculation :-)

10.97 The daily number of incoming calls $X$ is assumed to be a Poisson($\Lambda$) random variable, for some $\Lambda$ which itself is considered random following a gamma($50, 2$) distribution. The average daily number of incoming calls is to be estimated. This can be expressed as the parameter

$$\mu = E\Lambda(E(X|\Lambda)),$$

that is, the average over the $\Lambda$ distribution of the mean of $X$ in the model Poisson($\Lambda$). Use the mean of the prior distribution of $\Lambda$, the single observation $X$, and the mean of the posterior distribution of $\Lambda$ given $X$, respectively, to estimate $\mu$ in the cases (a), (b), and (c).

What estimate would you choose if you were the manager? Why?

11.1 We need to find $k$ such that

$$P(0 \leq \theta \leq kX) = 1 - \alpha,$$

where $X$ is exponential($\theta$). Since

$$P(0 \leq \theta \leq kX) = P(X \geq \theta/k) = \int_{\theta/k}^{\infty} (1/\theta) e^{-x/\theta} dx,$$

this is equivalent to finding $k$ such that

$$\frac{1}{\theta} \int_{\theta/k}^{\infty} e^{-x/\theta} dx = 1 - \alpha.$$

Now evaluate the integral and reduce this equation to

$$e^{-1/k} = 1 - \alpha.$$

Solution: Using $k = -\frac{1}{\log(1-\alpha)}$, the 100(1 - $\alpha$)% C.I. for $\theta$ is

$$\left(0, \frac{X}{-\log(1-\alpha)}\right).$$

Note: Since $1 - \alpha$ is less than 1, $-\log(1-\alpha)$ is a positive number. As $\alpha \to 0$, this number ↓ log 1 = 0 and the width of the interval ↑ ∞ (paying a price for higher confidence).

11.30 This problem involves a sample of size 10, and you can take the given numbers as the sample mean $\bar{X} = 5.68$ cm and the sample standard deviation $s = 0.29$ cm. Since a normal population (exactly) is
assumed, you are in the situation of wanting a 95% C.I. for $\mu$ in $N(\mu, \sigma^2)$ with $\sigma^2$ unknown. So you will use a C.I. of form

$$\bar{X} \pm k \frac{s}{\sqrt{n}}$$

with $k$ chosen using the $t$ distribution with $n - 1$ degrees of freedom. Complete the details :) .

11.34 Here you have a two-sample difference-of-means problem with sample sizes $n_1 = n_2 = 61$, respective sample means $\bar{X}_1 = 80.7$ minutes and $\bar{X}_2 = 88.1$ minutes, and respective sample standard deviations $s_1 = 19.4$ minutes and $s_2 = 18.8$ minutes. Since normality is not assumed, we do not use the $t$ distribution. Rather, we use the central limit theorem (CLT) to treat the sample means as approximately normal, and then use an interval of form

$$(\bar{X}_1 - \bar{X}_2) \pm k \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

with $k$ chosen using the standard normal distribution as an approximate 99% C.I. for $\mu_1 - \mu_2$. Complete the details :).

11.38 Obtain and justify the interval 0.68 ± 0.053, or (0.627, 0.733), for the proportion $\theta$.

11.49 Obtain and justify the interval $-0.288 \pm 0.084$, or $(-0.372, -0.204)$, for the difference of proportions $\theta_1 - \theta_2$.

11.53 I think that you should use the interval given in Theorem 11.9, page 367, with $n = 10$, $s^2 = (0.29)^2$, and $\alpha = 0.05$, to obtain a 95% C.I. for $\sigma^2$. What do you think? :)

11.57 Enjoy :) .

12.4 To use Table I, we apply the fact that if $X$ is binomial(20, $\theta$), then $20 - X$ is binomial(20, 1 - $\theta$). Then

$$\alpha = P(X \leq 16; X \text{ bin}(20, 0.90))$$
$$= P(20 - X \geq 4; X \text{ bin}(20, 0.90))$$
$$= P(Y \geq 4; Y \text{ bin}(20, 0.10))$$
$$= 1 - \sum_{k=0}^{3} P(Y = k; Y \text{ bin}(20, 0.10))$$
$$= 1 - (.1216 + \cdots + 0.1901) \quad \text{(via Table I)}$$
$$= 0.1329.$$

Similarly, obtain $\beta = 0.0159$.

12.10 Follow similar steps as in the text and in class.

12.15 Show that the MP test reduces to a critical region of form $C = \{\sum_{i=1}^{n} X_i^2 \geq K\}$ for some constant $K$. To determine $K$ for desired significance level $\alpha$, use the fact that if $X$ is $N(0, \sigma^2)$, then $X^2/\sigma^2$ is chi-square with 1 degree of freedom. Hence, by independence, under the null hypothesis $(\sum_{i=1}^{n} X_i^2)/\sigma_0^2$ is chi-square with $n$ degrees of freedom. Complete the details.
12.25 The data consists of independent samples of sizes $n_i$ from $N(\mu_i, \sigma_i^2)$, respectively, $i = 1, \ldots, k$. All parameters are unknown. Let $n = n_1 + \cdots + n_k$.

(a, part i) Under the null hypothesis that $\sigma_1^2 = \cdots = \sigma_k^2 = \sigma^2$, say, the likelihood function is given by

$$L(\mu_1, \ldots, \mu_k, \sigma^2) = \prod_{i=1}^k \left[ \left( \frac{1}{\sigma \sqrt{2\pi}} \right)^{n_i} e^{-\frac{1}{2\sigma^2} \sum_{j=1}^{n_i} (x_{ij} - \mu_i)^2} \right]$$

$$= \left( \frac{1}{\sigma \sqrt{2\pi}} \right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \mu_i)^2}$$

Now proceed as in Example 10.17, pp. 340–341. Take the log of $L$ and then take partial derivatives of $\log L$ with respect to $\mu_1, \ldots, \mu_k$ and $\sigma^2$. Set these all to zero to obtain $k + 1$ equations. For $\mu_i$, we obtain

$$\frac{\partial \log L(\mu_1, \ldots, \mu_k, \sigma^2)}{\partial \mu_i} = \frac{1}{\sigma^2} \sum_{j=1}^{n_i} (x_{ij} - \mu_i) = 0,$$

$i = 1, \ldots, k$. I leave it to you to derive the remaining equation and to show that these equations yield the maximum likelihood estimators stated in the problem. Also, for use in part (b), substitute these estimators for the parameters in the above likelihood and obtain that the maximum of the likelihood under the null hypothesis is given by

$$\left( \frac{1}{\sum_{i=1}^k \frac{(n_i-1)s_i^2}{n}} \right)^{1/2} \sqrt{2\pi} \frac{n}{n_i}.$$

(a, part ii) For the alternative that the variances are not all equal, the unrestricted likelihood is given by

$$L(\mu_1, \ldots, \mu_k, \sigma_1^2, \ldots, \sigma_k^2) = \prod_{i=1}^k \left[ \left( \frac{1}{\sigma_i \sqrt{2\pi}} \right)^{n_i} e^{-\frac{1}{2\sigma_i^2} \sum_{j=1}^{n_i} (x_{ij} - \mu_i)^2} \right]$$

$$= \left[ \prod_{i=1}^k \left( \frac{1}{\sigma_i \sqrt{2\pi}} \right) \right]^n e^{-\sum_{i=1}^k \frac{1}{2\sigma_i^2} \sum_{j=1}^{n_i} (x_{ij} - \mu_i)^2}$$

Again take logs, then partial derivatives, set equal to zero, and obtain the estimators stated in the problem. Also, for use in part (b), substitute these estimators for the parameters in the unrestricted likelihood and obtain that the maximum of the likelihood under the unrestricted model is given by

$$\prod_{i=1}^k \left( \frac{1}{\left( \frac{(n_i-1)s_i^2}{n} \right)^{1/2} \sqrt{2\pi}} \right)^{n_i}.$$

(b) This is now straightforward using the results obtained in (a).
12.26 Hint: divide numerator and denominator by $s_1^{n_1+n_2}$. Enjoy :)

12.41 Enjoy :)

12.44 Here, the statement of the problem should cite Exercise 12.21 instead of Exercise 12.32. For the present exercise, you may use the result of Exercise 12.21 (a) that the likelihood ratio test statistic is

$$
\lambda = \left( \frac{\bar{x}}{\theta_0} \right)^n e^{-(n\bar{x}/\theta_0) + n}.
$$

Now use the approximation to the null hypothesis distribution of log $\lambda$ provided by Theorem 12.2, namely chi-square with 1 degree of freedom. To make a decision, compare the test statistic value 7.86 with the cut-off point $\chi^2_{0.05, 1} = 3.84$.

13.2 The critical region for a level $\alpha$ MP test of this one-sided $H_0$ was found in Example 12.4 to be of the form:

$$
C = \{ (x_1, \ldots, x_n) : \bar{x} \geq \mu_0 + z_{\alpha} \frac{\sigma}{\sqrt{n}} \}.
$$

The Type II error probability then equals $\beta$, therefore, if

$$
P(C \mid \mu = \mu_1) = P \left( \bar{X} \geq \mu_0 + z_{\alpha} \frac{\sigma}{\sqrt{n}} \mid \mu = \mu_1 \right)
$$

$$
= P \left( \frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} \geq \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + z_{\alpha} \right)
$$

$$
= P \left( N(0, 1) \geq \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + z_{\alpha} \right)
$$

$$
= \beta.
$$

But also we must have

$$
P(N(0, 1) \geq z_{\beta}) = \beta.
$$

Comparison of these equations shows that we must have, therefore,

$$
\frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + z_{\alpha} = z_{\beta}.
$$

Now solve this equation for $n$.

13.20 Use the fact that the null hypothesis is to be rejected if and only if the $P$-value is less than the nominal level of significance.

13.26 For this problem, assume a sample of size $n = 45$ from $N(\mu, (8.6)^2)$. Test $H_0 : \mu = 84.3$ against the alternatives $\mu > 84.3$. Base the test on the observed sample mean $\bar{X} = 87.8$, and get the one-sided $P$-value: 0.0032.

13.35 Enjoy :)

13.47 Analogous to Example 13.6.

13.63 Analogous to Example 13.9.

13.74 Analogous to Example 13.10.
13.75 This also is analogous to Example 13.10. This problem also can be carried out in the style of Example 13.11. Although that example concerns a test of independence, and this problem concerns a test of homogeneity of distributions, the calculations and test procedures are the same, as discussed in class. Do the problem both ways, as a check. You should end up with a chi-square statistic of $\chi^2 = 8.03$. With reference to the appropriate chi-square distribution, find the $P$-value.

13.77 This also is analogous to Example 13.11.

13.78 This also is analogous to Example 13.11, except here homogeneity of distributions is being tested.

13.82 This is analogous to Example 13.12. Here the random variable $X$ is the number of cakes sold each day, out of a batch of three large chocolate cakes made for sale that day. We want to test goodness of fit to a binomial(3, $\theta$) model, i.e., to test whether $X$ has such a distribution for some (unknown) $\theta$, where $\theta$ is the probability that a given cake is sold. The data set consists of the observed sales for $n$ days. Since the outcome $X = 0$ occurred on 1 day, the outcome $X = 1$ on 16 days, the outcome $X = 2$ on 55 days, and the outcome $X = 3$ on 228 days, the total number of observations is $n = 1 + 16 = 55 + 228 = 300$.

We need to estimate the unknown probability $\theta$ that a cake becomes sold. Since the total number of cakes is $3 \times 300 = 900$ and the number of these that were sold is $0 \cdot 1 + 1 \cdot 16 + 2 \cdot 55 + 3 \cdot 228 = 810$, we have

$$\hat{\theta} = \frac{810}{900} = 0.9.$$ For the binomial(3, 0.9) model, the probabilities attached to 0, 1, 2, 3, respectively, are (via Table I) 0.001, 0.027, 0.243, 0.729. Hence the expected frequencies are given by these probabilities multiplied by 300: $e_1 = 0.3$, $e_2 = 8.1$, $e_3 = 72.9$, $e_4 = 218.7$. Since the expected frequency $e_1$ is less than 5, we combine this category with the adjacent one: $e_1 + e_2 = 8.4$ and combine the corresponding observed frequencies as well: $f_1 + f_2 = 1 + 16 = 17$. Now construct the test statistic

$$\chi^2 = \frac{(17 - 8.4)^2}{8.4} + \frac{(55 - 72.9)^2}{72.9} + \frac{(228 - 218.7)^2}{218.7} = 8.80 + 4.40 + 0.40 = 13.6.$$ The reference distribution for this statistic under $H_0$ is chi-square with degrees of freedom $m - 1 - t$, where $m$ is the number of count categories (after combining) and $t$ is the number of estimated parameters. In this case, therefore, the degrees of freedom is $3 - 1 - 1 = 1$. Now compute the $P$-value for the observed value 13.6 in the $\chi^2$ distribution. Via Table V, we obtain $P < 0.005$, i.e., $P$ is much smaller than 0.005. Conclusion: The data does not support the hypothesis of a binomial model.

Question: how would the results change if the baker made cherry pies instead of chocolate cakes?

14.7 First get the marginals,

$$g(x) = \int f(x, y) \, dy = \int_0^x 2 \, dy = 2x, \quad 0 < x < 1,$$
and

$$h(y) = \int_y^1 2 \, dx = 2(1 - y), \quad 0 < y < 1.$$
Then get the conditional distributions,

\[ w(y|x) = \frac{f(x, y)}{g(x)} = \frac{2}{2x} = \frac{1}{x}, \quad 0 < y < x, \]

and

\[ w(x|y) = \frac{1}{1 - y}, \quad y < x < 1. \]

Now proceed with parts (a) and (b) using routine integration. Find \( \mu_{Y|x} = x/2 \), \( \mu_{X|y} = (1 + y)/2 \), and \( E(X^mY^n) = 2/(n + 1)(m + n + 2) \). Part (c) is then straightforward.

14.16 Use straightforward partial differentiation of the appropriate least squares objective function,

\[ q = \sum_{i=1}^{n} [Y_i - (\hat{\alpha} + \hat{\beta}X_i + \hat{\gamma}X_i^2)]^2, \]

to obtain the equations

\[
\begin{align*}
\sum Y_i &= \hat{\alpha} \cdot n + \hat{\beta} \sum X_i + \hat{\gamma} \sum X_i^2 \\
\sum X_i Y_i &= \hat{\alpha} \cdot \sum X_i + \hat{\beta} \sum X_i^2 + \hat{\gamma} \sum X_i^3 \\
\sum X_i^2 Y_i &= \hat{\alpha} \cdot \sum X_i^2 + \hat{\beta} \sum X_i^3 + \hat{\gamma} \sum X_i^4.
\end{align*}
\]

14.17 Straightforward algebra :-)

14.21 Straightforward. Note that this result does not depend on any assumption of a special bivariate distribution for \((X, Y)\).


14.77 Follow EXAMPLES 14.9 and 14.10. For (b), the case of a shipment weighing 2400 pounds being moved 1200 miles, you should get a predicted average damage of $101.41.

14.83 Apply Exercise 14.16 above to get the relevant equations. Solve them as in Exercise 14.77 and apply the solution similarly. For the case when 6.5 grams of the chemical is added, you should get a predicted average drying time of 5.95 hours.