1. **Test Statistics:** Suppose you have a random sample of size n from a $N(\mu, \sigma^2)$ PDF. The Neyman-Pearson Lemma gives the following 6 statistics (or slight variations thereof) as the best use of the data under various assumptions:

$$(a)\ \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \qquad (b)\ \frac{\bar{X}-\mu_0}{S/\sqrt{n}} \qquad (c)\ \frac{(n-1)\,S^2}{\sigma_0^2}$$

$$(d)\ \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \qquad (e)\ \frac{\bar{X}-\mu}{S/\sqrt{n}} \qquad (f)\ \frac{(n-1)\,S^2}{\sigma^2}\,.$$

   Explain when and how you should use these statistics appropriately for testing and reporting confidence intervals.

   Hint: State the purpose and assumptions about the normal parameters.

2. **Power-Like Calculation:** The Texas Department of Agriculture is responsible for inspecting gas station pumps to ensure their performance is within *acceptable tolerances.* Section 13.024 of the Texas statute defines the standard for liquid capacity to be (a) a gallon; (b) a barrel (31.5 gallons); or (c) a hogshead (2 barrels). NIST (National Institute for Standards and Technology) defines the basic tolerance for a gas pump that has been in service for more than 30 days to be one cubic inch, plus one cubic inch per each gallon indicated.

   Thus if we draw a 5-gallon test draft, the tolerance will be ± 6 cubic inches, which converts to ± 0.026 gallons. The test error on a single trial must not exceed this amount to certified. (Dispensing either too little or too much is considered a reason not to certify.)

   Consider a simple design of a gas pump that uses an impeller mechanism, dispensing a "quantum" amount of gas described by a Normal random variable X with mean $\mu_X = 0.01$ gallon and standard deviation $\sigma_X = 0.0003$ gallon. Each such amount of gas dispensed is independent of the previous amount. Then the "actual" amount pumped when the gas pump reads 5 gallons can be modeled after 500 "clicks" as

$$Y = \sum_{i=1}^{500} X_i\,.$$

   (a) Justify this formula for a 5-gallon test. What is the PDF of Y?

   Hint: Your answer should include μ_Y and σ_Y as well as the form of the PDF.

   (b) There is no adjustment possible for σ_X on the pump. It is fixed for an impeller pump. However, there is an adjustment for μ_X. What is the formula for the probability that the pump is certified, as μ_X is adjusted from 0.0099 to 0.0101? Graph it.

   (c) If you wish to ensure that your pump is certified with probability at least 95% , what interval should μ_X fall in? Mark these points on your graph.

   Hint: An approximate answer from your graph is sufficient.

3. **Goodness-of-Fit Test:** How good is the Uniform random generator provided with **R**? To really stress it, let us subject it to a rigorous Pearson goodness-of-fit test. Generate a large sample, $\{U_1, U_2, \ldots, U_n\} \sim Unif(0,1)$ of length $n = 10^5$, which you will count into 1000 bins,

using the following commands:

> `set.seed(8736); n=1e5; nb=1000`	(sample size and number of bins)
> `tk = seq(0,1, , nb+1)`	(bin boundaries; notice the extra comma)
> `nk = hist( runif(n), tk )$counts`	(`nk` contains the 1000 bin counts)

You can check your bin counts: `105 113 97 ... 104 102 111`.

(a) Perform a goodness-of-fit test at the 5% level of the hypotheses

$$H_0 \,:\, U_i \sim Unif(0,1) \quad versus \quad H_1\,;\, U_i \not\sim Unif(0,1)$$

Hint: Give the critical region C, the test statistic, and your decision.

(b) What is the p-value for your experiment? Sketch it.

4. **Pooled Variance:** Consider the two-sample t-test setup: $\{X_1, X_2, \ldots, X_{n_x}\}$ and $\{Y_1, Y_2, \ldots, Y_{n_y}\}$ are 2 random samples from $N(\mu_x, \sigma^2)$ and $N(\mu_y, \sigma^2)$, with all parameters unknown, but with $\sigma_x^2 = \sigma_y^2 = \sigma^2$. We wish to test

$$H_0 \,:\, \mu_x = \mu_y \qquad versus \qquad H_1\,;\, \mu_x \neq \mu_y\,.$$

(a) Show that both

$$S_X^2 = \frac{1}{n_x - 1}\sum_{i=1}^{n_x}(X_i - \bar{X})^2 \quad \text{and} \quad S_Y^2 = \frac{1}{n_y - 1}\sum_{i=1}^{n_y}(Y_i - \bar{Y})^2$$

are both unbiased for σ^2. *Hint: Just cite the appropriate result in our book.*

(b) Conclude that the pooled variance estimator

$$S_P^2 = \frac{(n_x - 1)S_X^2 + (n_y - 1)S_Y^2}{n_x + n_y - 2}$$

is also unbiased for σ^2.
Hint: Consider using the MGF technique.

(c) Compare the variances of S_X^2, S_Y^2, and and S_P^2. Which is best?
Hint: Recall the first two moments of a χ_p^2 PDF are p and $2p$.

5. **Test of Independence:** One often hears stories of elderly individuals dying shortly after significant events, such as birthdays, anniversaries, or holidays. A study was conducted in California for the years 1960-1984 examining the mortality patterns of elderly Chinese women who died of natural causes immediately before and after the Chinese Harvest Moon Festival, for which the senior woman of the household plays a central ceremonial role. Their data were compared to elderly Jewish women, and are given in the following table:

Time of Death	Chinese	Jewish	Total
2^{nd} week before festival	55	141	196
1^{st} week before festival	33	145	178
1^{st} week after festival	70	139	209
2^{nd} week after festival	49	161	210
Total	207	586	793

(a) Set up an appropriate Pearson χ^2 hypothesis test of the independence of the rows and columns characteristics, giving the null and alternative hypotheses. What is the critical region at the $\alpha = 5\%$ level?

(b) What is the test statistic and your decision? Write down the matrix of expectations.

(c) What is the p-value of your test?

(d) Explain how the pattern in the table relates to your decision.

6. **Best Critical Region:** Suppose you have a random sample $\boldsymbol{X} = \{X_1, X_2, \ldots, X_n\}$ from the geometric PMF

$$\boxed{p_X(x) = p \cdot (1-p)^{x-1}\,,\; x = 1, 2, \ldots} \qquad \mu_X = \frac{1}{p} \text{ and } \sigma_X^2 = \frac{1-p}{p}\,.$$

We wish to test

$$H_0 \,:\, p = p_0 \qquad versus \qquad H_1\,;\, p = p_1 < p_0\,.$$

(Thus we are wondering if the time to the first success is longer than thought.)

(a) Find the log-likelihood $\ell(p|\boldsymbol{x})$. Then use it in the Neyman-Pearson Lemma to find the generic form of the best critical region, C.

Hint: You may assume $\bar{x} > 1$. Why?

(b) Suppose $n = 12$ and

$$p_0 = \frac{1}{4} \quad \text{and} \quad p_1 = \frac{1}{5}\,.$$

Find the best critical region at the level $\alpha = 5\%$.

Hint: Use the CLT.

(c) What is the power of your test at the alternative hypothesis?

7. **Equivalence of Tests and Confidence Intervals:** Suppose you have a random sample of size n from a Normal PDF with unknown mean and variance. You wish to test the null hypothesis $H_0 : \mu = \mu_0$ versus a two-sided alternative. Show that you reject the null hypothesis if-and-only-if the corresponding confidence interval for μ does not contain the μ_0.

Hint: Your answer should be both algebraic and graphical.

8. **ANOVA:** A small brewery decided to experiment with label designs on its summer ale. Twenty stores of similar size were selected. The number of cases sold over a two-month period were recorded:

Design						Total
1	11	17	16	14	15	73
2	12	10	15	19	11	67
3	23	20	18	17	21	99
4	27	33	22	26	28	136

   (a) Test the null hypotheses that the mean number of sales at a store are the same for all four label designs, versus the alternative that some are different. Use $\alpha = 1\%$.

   (b) What is the p-value?

9. **Comparing Correlations:** In addition to the Father (F) – Son (S) height data discussed in Chapter 1, Pearson and Lee also collected heights of Mothers (M) and Daughters (D). The sample correlations (which can be denoted by either $\hat{\rho}$ or R) are

$$\hat{\rho}_{FD} = 0.511 \qquad n_{FD} = 1376$$
$$\hat{\rho}_{MS} = 0.494 \qquad n_{MS} = 1057\,.$$

The subscripts indicate the Father-Daughter and Mother-Son correlations, and the sample sizes upon which they are based. You may assume these correlation estimates are independent since based upon different families. (Recall the Father-Son correlation was also about 0.5.)

   (a) Devise a level $\alpha = 5\%$ test for the hypotheses

$$H_0 : \rho_{FD} = \rho_{MS} \quad \text{versus} \quad H_1 : \rho_{FD} \neq \rho_{MS}\,,$$

   What is the value of your test statistic? Your decision?
   Hint: This requires you putting several pieces (facts) together in a manner similar to some of the tests we have studied. Not a formal Neyman-Pearson procedure. Show your work.

   (b) What is the p-value for these data? Explain how your answer is consistent with part (a). *Hint: Highlight the area representing the p-value.*

10. **Linear Regression:** Using the data $\{(x_i, y_i),\ i = 1, \ldots, 24\}$ from the 1974–1997 Boston

Marathon winning times (in minutes) for the women runners ($n = 24$), we summarize

$$\sum_{i=1}^{24} x_i = 47{,}652 \qquad \sum_{i=1}^{24} (x_i - \bar{x})^2 = 1150$$
$$\sum_{i=1}^{24} y_i = 3620.867 \qquad \sum_{i=1}^{24} (x_i - \bar{x})\, y_i = -1145.45$$

   (a) Find the best fitting least-squares straight line and add it to the scatter diagram provided below. Give the constants $\hat{a}$ and $\hat{b}$, using the form given in Equation (8.26).

   (b) Some of the winning times for years since 1997 are given in the table:

Year	2000	2003	2006	2009	2012	2015	2018
Time	146.2	145.3	143.6	152.3	151.8	144.9	159.9

We wish to see how well our linear equation predicted the "future." Compute $\hat{\sigma}_\epsilon^2$ and sketch 95% population intervals for the years 1980–2020. Add the data in the table to the scatter diagram. How far into the future do the predictions seem to hold? Comments?

(c) What is your prediction for 2019 (run on April 15)? How many standard deviations was the actual winning time of 143.48 minutes from your prediction?

Hint: Give the value of T_{n-1}.

Hints: In part (b), give the form of the intervals you have plotted. Do not worry about adjusting the α*-level due to multiple testing (i.e., no Bonferroni adjustment to* α*).*