1. (10 points) Suppose we have $n = 3$ random samples from a PDF with mean equal to 0 and finite variance $\sigma_X^2 > 0$. We introduce 3 new random variables derived from $X_1, X_2,$ and X_3:

$$\begin{aligned} Y_1 &= X_1 \\ Y_2 &= X_1 + X_2 \\ Y_3 &= X_1 + X_2 + X_3 \end{aligned}$$

   What are the three **correlations**, ρ_{ij}, between Y_i and Y_j?

2. (10 points) Suppose the random variables X and Y are standard bivariate Normal with correlation coefficient ρ. What is the exact PDF of the random variable

$$V = 3X - Y \ \ ?$$

   Hint: $$E\left[e^{tV}\right] = \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} e^{t(3x-y)} \cdot \frac{1}{2\pi\sqrt{1-\rho^2}} \, e^{\frac{-(x^2-2\rho xy+y^2)}{2(1-\rho^2)}} = e^{(5-3\rho)t^2} \, .$$

3. (15 points) Suppose the random variable T follows a particular Negative Exponential PDF given by

$$f_T(t) = \lambda \, e^{-\lambda t} \quad \text{for } \lambda, t > 0 \, .$$

   Then the random variable

$$S_2 = T_1 + T_2$$

   represents the waiting time for the second occurrence of an event in a Poisson process.

   (a) (5 points) Given a random sample of size n, $\{u_i\} \sim Unif(0,1)$, how can you generate a random sample $\{t_i\}$ from f_T using the probability integral transformation?

   (b) (5 points) Find the exact PDF of the random variable S_2 using the moment-generating-function technique.

   (c) (5 points) For the particular value $\lambda = 2$, use your PIT formula in part (a) to generate a sample of size $n = 1{,}000$ for S_2. Plot a probability histogram of your sample and use the appropriate R function to overlay the true PDF curve.

4. (10 points) Suppose we have a random sample from the $Unif[-\theta, 2\theta]$, where $\theta > 0$.

   (a) (5 points) What is the MLE of θ?

   (b) (3 points) What is (are) the sufficient statistic(s) for θ?

   (c) (2 points) What is $\hat{\theta}_{MLE}$ for the random sample of size $n = 10$

$$(-0.686, 6.825, 1.135, 8.245, 9.107, -4.317, 2.922, 8.386, 3.272, 1.849) \, ,$$

which were generated by `> set.seed(123); round(runif(10,-5,10),3).`

5. (10 points) An entrepreneur is considering offering a comprehensive dental insurance plan for Texans. Given the plan design, she estimates that the average benefit would be \$1,000. There would be a large spread of benefits, which is represented by her estimate of \$4,500 for the standard deviation. In modeling the finances of the first year, she wishes to understand the effect if n individuals sign up on day 1. Law requires that the premiums (reserves) collected must cover the payouts with probability 99.99%. For n ranging from 10^2 to 10^6, compute the average premium *for each participant* that would exactly satisfy the reserves requirement. State what assumptions (and approximations) you are using to compute the 99.99 percentile? Graph the curve and see how close to the average benefit of \$1,000 it gets.

   Hints: (1) Assume each participant is charged the same premium; (2) We are ignoring all other expenses involved in administering the plan (just focusing on the actual dental service costs); (3) You may wish to invoke the CLT.

6. (10 points) Consider n nice, smooth, monotone increasing functions $\{u_i(\cdot), i = 1, \ldots, n\}$. Intuitively, if $X_1, X_2, \ldots, X_n$ is a random sample, then defining n new random variables

$$Y_1 = u_1(X_1), \quad Y_2 = u_2(X_2), \quad \ldots, \; Y_n = u_n(X_n)$$

   should result in a random sample as well *if all the u_i are identical.* After all, there is no "overlap" among the samples, and the same transformation is applied to each X_i.

   Here we wish to demonstrate the fact that Y_1 and Y_2 are independent (when u_1 and u_2 are not identical) for a sample of size $n = 2$, using our usual the bivariate change of variables notation (instead of u_1 and u_2):

$$\begin{aligned} y_1 &= y_1(x_1, x_2) = y_1(x_1) \\ y_2 &= y_2(x_1, x_2) = y_2(x_2)\,, \end{aligned}$$

   since the transformations involve only one of x_1 and x_2. Now since $u_1(\cdot)$ and $u_2(\cdot)$ are smooth, monotone transformations, the transformations are 1–1 and the inverse transformations exist:

$$\begin{aligned} x_1 &= x_1(y_1) \\ x_2 &= x_2(y_2)\,, \end{aligned}$$

   Since X_1 and X_2 are a random sample,

$$f_{X_1,X_2}(x_1, x_2) = f_X(x_1) f_X(x_2)\,.$$

   (a) (6 points) Use the bivariate change of variables technique to derive $g(y_1, y_2)$.

   (b) (4 points) Can you conclude that Y_1 and Y_2 are independent? Why?

7. (10 points) Show that if the random variables X and Y follow the standard Bivariate Normal PDF with single parameter ρ,

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{\frac{-(x^2-2\rho xy+y^2)}{2(1-\rho^2)}} \,,$$

Show that if $\rho \neq 0$, then X and Y are **not** independent.

Hint: One possible approach is to find the marginal distributions of X and Y and the conditional distribution of $Y|X=x$.

8. (15 points) Suppose we have two **unbiased estimators**, X and Y, for an unknown parameter θ. But the variance of Y is greater than the variance of X; hence, X is more efficient. Furthermore, X and Y are *not independent*, but are **correlated**. In fact,

$$\sigma_Y^2 = 4\,\sigma_X^2 \quad \text{and} \quad \rho_{XY} = -\frac{1}{2}\,.$$

Show the following three estimators — $\hat{\theta}_1$, $\hat{\theta}_2$, and $\hat{\theta}_3$ — are also unbiased for θ and compute their variance. Order the 3 estimators and determine which is "best?"

(a) (5 points)

$$\hat{\theta}_1 = X \quad \text{itself?}$$

(b) (5 points)

$$\hat{\theta}_2 = \frac{X+Y}{2}\,, \quad \text{their average?}$$

(c) (5 points)

$$\hat{\theta}_3 = \frac{5}{7}X + \frac{2}{7}Y\,, \quad \text{a weighted average?}$$

9. (15 points) In a 310 class with n students, each student was asked to spin a nickel 15 times and record the number of heads, resulting in the data vector $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$. Our interest is focused on p, the probability of obtaining a head on a single spin.

(a) (2 points) $\{X_1, X_2, \ldots, X_n\}$ is a random sample from what PMF?

(b) (2 points) What is the log-likelihood, $\ell(p|\boldsymbol{x})$, for the unknown parameter p?

(c) (2 points) What is the sufficient statistic for p?

(d) (3 points) What is the MLE, $\hat{p}_{MLE}$ of p? Hint: Don't need to check 2nd derivative.

(e) (3 points) What is the expectation of $\hat{p}_{MLE}$?

(f) (3 points) What is the variance of $\hat{p}_{MLE}$? Hint: Careful....

10. (5 points) Show

$$E\left[I(X \in A)\right] = Pr(X \in A)\,.$$