MTH211A: Theory of Statistics

Problem set 1

1. Let the random variable X have pdf
$$f(x) = \sqrt{\frac{2}{\pi}} \exp\{-x^2/2\}, \qquad x > 0.$$
   (a) Find $E(X)$ and $\operatorname{var}(X)$.
   (b) Find an appropriate transformation $Y = g(X)$ and α, β, so that $Y \sim \texttt{Gamma}(\alpha, \beta)$.

2. Let X is distributed as $\texttt{Gamma}(\alpha, \beta)$ distribution, $\alpha, \beta > 0$. Then show that the r-th order population moment
$$E\left(X^r\right) = \beta^{-r}\frac{\Gamma(r+\alpha)}{\Gamma(\alpha)}, \qquad r > -\alpha.$$

3. Let the bivariate random variable (X, Y) has a joint pdf
$$f_{X,Y}(x, y) = \begin{cases} C(x+2y) & \text{if} \quad 0 < y < 1,\ 0 < x < 2, \\ 0 & \text{otherwise.} \end{cases}$$
   (a) Find the marginal distribution of Y.
   (b) Find the conditional distribution of Y given $X = 1$.
   (c) Compare the expectations of the above two distributions of Y.
   (d) Find the covariance between X and Y.
   (e) Find the distribution of $Z = 9/(2Y + 1)^2$.
   (f) What is $P(X > Y)$?

4. Let $X \sim \texttt{normal}(0, 1)$. Define $Y = -X\mathbb{I}(|X| \leq 1) + X\mathbb{I}(|X| > 1)$.
   (a) Find the distribution of Y (Hint: Start by finding the CDF of Y).
   (b) Prove or disprove: The distribution of (X, Y) is bivariate normal.
   (Hint: Argue that the distribution of $X + Y$ is not continuous.)

5. Let $X \sim \texttt{normal}(0, 1)$. Define $Y = \operatorname{sign}(X)$ and $Z = |X|$. Here $\operatorname{sign}(\cdot)$ is a $\mathbb{R} \to \{0, 1\}$ function such that $\operatorname{sign}(a) = 1$ if $a \geq 0$, and $\operatorname{sign}(a) = -1$ otherwise.
   (a) Find the marginal distributions of Y and Z.
   (b) Find the joint CDF of (Y, Z). Hence or otherwise prove that Y and Z are independently distributed.

6. Suppose the distribution of Y, conditional on $X = x_0$ is $\texttt{normal}(x_0, x_0^2)$, and the marginal distribution of X is $\texttt{uniform}(0, 1)$. Show that $Z = Y/X$ and X are independently distributed. Find the distribution of Y/X.

7. (a) Let (X, Y) is jointly distributed as $\mathtt{N}_2(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$. Suppose (X, Y) are uncorrelated in the sense that $\operatorname{cov}(X, Y) = E\left[\{X - E(X)\}\{Y - E(Y)\}\right] = 0$. Then show that X and Y are independently distributed.

   (b) Let $\mathbf{X} = (X_1, \cdots, X_k)^\top$ is distributed as a k-variate normal distribution with parameters $\boldsymbol{\mu}$ and $\Sigma = \operatorname{Diag}\left(\sigma_1^2, \cdots, \sigma_k^2\right)$, $\sigma_j > 0$ for all $j = 1, \ldots, k$. Show that $X_1, \cdots, X_k$ are mutually independent.

8. Let $X_i \overset{\text{IID}}{\sim} \text{normal}\left(\mu, \sigma^2\right)$. Define $U = (a_1 X_1 + \cdots + a_n X_n)$ and $V = (b_1 X_1 + \cdots + b_n X_n)$.

   (a) Show that (U, V) jointly follow a bi-variate normal distribution. Identify the parameters of the distribution. (Hint: For any $\mathbf{c} \in \mathbb{R}^2$, show that $\mathbf{c}^\top \begin{bmatrix} U \\ V \end{bmatrix}$ has an univariate normal distribution.)

   (b) Find conditions on $\mathbf{a}$ and $\mathbf{b}$ such that U and V are independently distributed.
   (Hint: Apply problem 7.)

9. Suppose $X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathtt{normal}\left(\mu_x, \sigma^2\right)$, $Y_1, \cdots, Y_m \overset{\text{IID}}{\sim} \mathtt{normal}\left(\mu_y, \sigma^2\right)$, and all the random variables $\{X_1, \cdots, X_n, Y_1, \cdots, Y_m\}$ are mutually independent. Then find the distribution of $T := S_X^{\star 2} / S_Y^{\star 2}$, where $S_X^{\star 2}$ and $S_Y^{\star 2}$ are the unbiased sample variances of X and Y, respectively.

10. Let $X_1, \cdots, X_n$ be iid random variables with continuous CDF F_X, and suppose $E(X_1) = \mu$. Define the random variables $Y_1, \cdots, Y_n$ as follows:

$$Y_i = \begin{cases} 1 & \text{if} X_i > \mu \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Find $E(Y_1)$.

   (b) Find the distribution of $\sum_{i=1}^n Y_i$.

11. Let $X_1, \cdots, X_n \overset{\text{IID}}{\sim} \mathtt{normal}\left(\mu, \sigma^2\right)$, and S_n^2 be the sample variance. Find a function of S_n^2, say $g\left(S_n^2\right)$, which satisfies $E\left[g\left(S_n^2\right)\right] = \sigma$. (Hint: You may use problem 2.)

12. Let $\mathbf{X}_1, \cdots, \mathbf{X}_n$ be a random sample from a k-variate distribution F, where $\mathbf{X}_i = (Y_{1,i}, \cdots, Y_{k,i})^\top$. Let the expectation $E\left(\mathbf{X}_1\right) = \boldsymbol{\mu}$ and the variance covariance matrix of F, $\Sigma = E\left\{\left(\mathbf{X}_1 - \boldsymbol{\mu}\right)\left(\mathbf{X}_1 - \boldsymbol{\mu}\right)^\top\right\}$, have all finite components, and $S_n = n^{-1} \sum_{i=1}^n \left(\mathbf{X}_i - \bar{\mathbf{X}}_n\right)\left(\mathbf{X}_i - \bar{\mathbf{X}}_n\right)^\top$ be the sample variance-covariance matrix, where $\bar{\mathbf{X}}_n = n^{-1} \sum_{i=1}^n \mathbf{X}_i$. Show that

   (a) $\bar{\mathbf{X}}_n = \left(\bar{Y}_1, \cdots, \bar{Y}_k\right)^\top$

   (b) For all $j, l = 1, \ldots, k$, the (j, l)-th component of Σ is $\sigma_{j,l} = E\left[\{Y_j - E(Y_j)\}\{Y_l - E(Y_l)\}\right]$. Also, for all $j, l = 1, \ldots, k$, the (j, l)-th component of S_n is $S_{j,l} = n^{-1} \sum_{i=1}^n \left(Y_{j,i} - \bar{Y}_j\right)\left(Y_{l,i} - \bar{Y}_l\right)$.

   (c) $E\left(\bar{\mathbf{X}}_n\right) = \boldsymbol{\mu}$ and $\operatorname{var}\left(\bar{\mathbf{X}}_n\right) = E\left\{\left(\bar{\mathbf{X}}_n - \boldsymbol{\mu}\right)\left(\bar{\mathbf{X}}_n - \boldsymbol{\mu}\right)^\top\right\} = n^{-1}\Sigma$.

   (d) $E\left(nS_n\right) = (n-1)\Sigma$.

13. Let $X_1, \cdots, X_n$ be iid with pdf f_X and CDF F_X, and let $r < s$. Then find the conditional distribution of $X_{(r)}$ given $X_{(s)} = y$ in terms of f_X and F_X. In particular, if $X_1, \cdots, X_n$ is a random sample from $\mathtt{uniform}(0, 1)$ distribution, then can you identify the conditional distribution of $X_{(r)}$ given $X_{(s)} = y$?

14. Let $X_1, \cdots, X_n$ be iid with pdf f_X and CDF F_X. Find the CDF of r-th order statistics $X_{(r)}$. Hence derive the pdf of $X_{(r)}$.

15. Let Y have a $\mathtt{Cauchy}(0, 1)$ distribution.

   (a) Find the CDF of Y.

   (b) Hence provide a method of simulating random samples from $\mathtt{Cauchy}(0, 1)$ distribution, starting from $\mathtt{uniform}(0, 1)$ random variables.