Stat 111 Week 3: Joint and Conditional Distributions

1. **Joint, Marginal and Conditional Densities** (Rice 3.1-3.5, Blitzstein 7.1)

   a) Review the concept of a joint pdf and describe the corresponding joint cdf. Review how you find marginal and conditional distributions from a joint distribution.

   b) As an example, suppose X and Y have joint pdf $f_{xy}(x, y) = \lambda^2 e^{-\lambda y} I_{(0<x<y)}$. Show that the marginal distributions are $X \sim$ Exponential(λ) and $Y \sim$ Gamma($2, \lambda$).

   c) Show that $Y|X = x$ is a translated Exponential, and $X|Y = y$ is Uniform.

   d) Review the factorization theorem for independence. Show that X and Y are not independent, but that X and $Y - X$ are independent.

   e) A bivariate differential argument approximates $P(X \in [x, x+dx), Y \in [y, y+dy))$ by $f_{xy}(x, y)dxdy$. Use this approach to show that this is the joint pdf for the time of the first event (x) and the time of the second event (y) for a Poisson process with rate λ events per unit time. Show this agrees with what you find setting $X = X_1$ and $Y = X_1 + X_2$, for $X_1, X_2 \overset{\text{i.i.d.}}{\sim}$ Exponential(λ).

2. **Covariance and Correlation** (Rice Section 4.3, Blitzstein 7.3)

   a) Define the covariance σ_{xy} and the correlation ρ_{xy} for random variables X and Y and derive the computation formula Cov(X,Y)=$E(XY) - E(X)E(Y)$. Compare to formulas for Var(X).

   b) As an example, consider a fat coin that has some probability of landing on its edge. Suppose the probability of it landing heads is p and the probability of it landing tails is also p ($0 < p \leq 0.5$). If you flip the coin $n = 2$ times and let X be the number of heads and Y the number of tails, find the covariance and correlation of X and Y as a function of p, and give the values when $p = 0.4$.

   c) Derive the formulas for Var($X + Y$) and Var($X - Y$). Use the fact that $Var(X + Y)$ and $Var(X - Y)$ are both non-negative to prove that $-1 \leq \rho_{xy} \leq 1$ (proving either the upper or lower bound will be sufficient for the presentation).

   d) Show that independent implies uncorrelated, and explain why uncorrelated does not always imply independent. As an example, consider X and Y uniformly distributed over the triangular region $0 < |Y| < X < 1$.

3. **Vector Random Variables** (Rice 14.4.1, Blitzstein 7.5)

   a) Introduce the concept of an n-dimensional random vector $\mathbf{X}$, and the generalizations of the mean and variance to the $n \times 1$ mean vector $\boldsymbol{\mu}$ and $n \times n$ covariance matrix $\mathbf{V}$ (Rice, 14.4.1). Show how the usual definition of Cov(X_i, X_j) arises as the expectation of the i, jth element (and the j, ith element) of the random $n \times n$ matrix $(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^{\mathrm{T}}$. Give $\boldsymbol{\mu}$ and $\mathbf{V}$ for an $n \times 1$ vector of independent random variables with means μ_i and variances σ_i^2, for $i = 1, \ldots, n$.

   b) Suppose $\mathbf{X}$ is $n \times 1$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{V}$. Explain the formulas for the mean vector and covariance matrix for linear transformations:

   $$E(\mathbf{AX} + \mathbf{C}) = \mathbf{A}E(\mathbf{X}) + \mathbf{C}, \qquad \mathrm{Var}(\mathbf{AX} + \mathbf{C}) = \mathbf{AXA}^{\mathrm{T}}$$

   where $\mathbf{Y} = \mathbf{AX} + \mathbf{C}$, for $\mathbf{A}$ an $m \times n$ and $\mathbf{C}$ an $m \times 1$ matrix of constants.

c) Show that

$$\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \left(\sum_{i=1}^{n} \text{Var}(X_i)\right) + 2\left(\sum_{i<j} \text{Cov}(X_i, X_j)\right).$$

which is the sum of all of the elements in the covariance matrix $\mathbf{V}$.

d) Show that $\text{Cov}(aX + b, c + dY) = ad\text{Cov}(X, Y)$ and $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$

e) For $X_1, \ldots, X_n$ iid random variables, show that $\bar{X}$ is uncorrelated with $X_i - \bar{X}$, for any $i = 1, \ldots, n$. Use the vector matrix formulation, and also show a scalar argument.

4. **Multivariate Normal Variables** (Blitzstein 7.5)
Definition: A random vector $(X_1, \ldots, X_n)$ follows a *Multivariate Normal* (MVN) distribution if $t_1X_1 + \ldots + t_nX_n$ follows a Normal distribution for any choice of constants $t_1, \ldots, t_n$.

a) Explain why this implies that any sample of n *independent* Normal random variables follows an dimensional MVN distribution.

b) Explain why, if X_i is part of a MVN vector, then X_i is Normal. Also give an example of variables X and Y that have marginal Normal distributions, but that are not part of a multivariate Normal vector.

c) The n-dimensional multivariate Normal density with $n \times 1$ mean vector $\boldsymbol{\mu}$ and $n \times n$ covariance matrix $\mathbf{V}$ has a proper joint pdf if and only if $\mathbf{V}$ is positive definite, and hence invertible. We write $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \mathbf{V})$, and the joint pdf is

$$f(\mathbf{x}) = (2\pi)^{-n/2}|\mathbf{V}|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}}\mathbf{V}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right]$$

Show that this is a generalization of the univariate Normal density ($n = 1$) and that the vector of $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$ has a joint pdf of this form (what are $\boldsymbol{\mu}$ and $\mathbf{V}$?).

d) Let $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ represent a multivariate Normal vector of dimension $m + n$, and suppose $\text{Cov}(X_i, Y_j) = 0$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$. Show that the joint density of the X's and Y's factors into a joint pdf for X_i's and a joint pdf for the Y_j's, meaning these are independent random vectors. State the general result about correlation and independence for elements of a multivariate Normal vector.

5. **Bivariate Normal Variables**

a) Write out the multivariate Normal density in the $n = 2$ case, in terms of the scalar elements (don't try this for dimensions over 2). Show the two elements are independent if and only if $\rho = 0$.

b) Let Z_1 and Z_2 be independent $N(0, 1)$ variables. For some value $\rho \in (-1, 1)$, set $X = Z_1$ and $Y = \rho Z_1 + \sqrt{1 - \rho^2}Z_2$. Verify that the vector (X, Y) is bivariate Normal with correlation ρ.

c) Let $Z \sim N(0, 1)$ be independent of $S = \pm 1$ with probability $1/2$. Let $X = Z$ and $Y = SZ$. Show that X and Y are uncorrelated standard Normal variables, but are not independent. Argue that, to be a bivariate Normal vector, X and Y must be non-trivial linear combinations of $Z_1, Z_2 \overset{\text{i.i.d.}}{\sim} N(0, 1)$, meaning both Z_1 and Z_2 get some weight with either X or Y

d) For X and Y bivariate Normal, identify the conditional distribution of $Y|X = x$. Give the more general result when X and Y have arbitrary means and standard deviations, and note the connections with simple linear regression.