Joint p.d.f. and Independent Random Variables

Let $X$ and $Y$ be two discrete r.v.’s and let $R$ be the corresponding space of $X$ and $Y$. The joint p.d.f. of $X = x$ and $Y = y$, denoted by $f(x, y) = P(X = x, Y = y)$, has the following properties:

(a) $0 \leq f(x, y) \leq 1$, $f(x, y) \geq 0$ for $-\infty < x, y < \infty$.

(b) $\sum_{(x,y) \in R} f(x, y) = 1$, $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy = 1$.

(c) $P[(X,Y) \in A] = \sum_{(x,y) \in A} f(x, y) ( \int \int f(x, y) )$, $A \subset R$.

The marginal p.d.f. of $X$ is defined as $f_X(x) = \sum_y f(x, y)$ $(\int_{-\infty}^{\infty} f(x, y) \, dy)$, $x \in R_x$.

The marginal p.d.f. of $Y$ is defined as $f_Y(y) = \sum_x f(x, y)$ $(\int_{-\infty}^{\infty} f(x, y) \, dx)$, $y \in R_y$.

The random variables $X$ and $Y$ are independent iff $f(x, y) \equiv f_X(x)f_Y(y)$ for $x \in R_x$, $y \in R_y$.

Example 1. $f(x, y) = (x + y)/21$, $x = 1, 2, 3$; $y = 1, 2$, then $X$ and $Y$ are not independent.

Example 2. $f(x, y) = (xy^2)/30$, $x = 1, 2, 3$; $y = 1, 2$, then $X$ and $Y$ are independent.

The collection of $n$ independent and identically distributed random variables $X_1, X_2, \ldots, X_n$, is called a random sample of size $n$ from the common distribution, say, $X_j \sim N(0,1)$, $1 \leq j \leq n$. 
Sampling Distribution Theory

♣ The collection of \( n \) independent and identically distributed random variables \( X_1, X_2, \ldots, X_n \), is called a random sample of size \( n \) from the common distribution, e.g., \( X_j \sim N(0, 1) \), \( 1 \leq j \leq n \).

♣ Some functions of a random sample, called statistics, are of interest, for examples, mean and variance. Sampling distribution theory refers to the derivation of distributions for functions of a random sample.

**Theorem 1:** Let \( X_1, X_2, \ldots, X_n \) be \( n \) independent r.v.'s with respective means \( \{\mu_i\} \) and variances \( \{\sigma^2_i\} \), then \( Y = \sum_{i=1}^{n} a_i X_i \) has mean \( \mu_Y = \sum_{i=1}^{n} a_i \mu_i \) and variance \( \sigma^2_Y = \sum_{i=1}^{n} a_i^2 \sigma^2_i \), respectively.

**Theorem 2:** Let \( X_1, X_2, \ldots, X_n \) be \( n \) independent r.v.'s with respective moment-generating functions \( \{M_i(t)\} \), \( 1 \leq i \leq n \), then the moment-generating function of \( Y = \sum_{i=1}^{n} a_i X_i \) is \( M_Y(t) = \prod_{i=1}^{n} M_i(a_i t) \).

**Corollary:** If \( X_1, X_2, \ldots, X_n \) are observations of a random sample from a distribution with moment-generating function \( M(t) \), then

(a) \( M_Y(t) = \prod_{i=1}^{n} M(t) \), where \( Y = \sum_{i=1}^{n} X_i \).

(b) \( M_{\bar{X}}(t) = \prod_{i=1}^{n} M(t/n) \), where \( \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \).

**Example 1:** Let \( X_i \sim b(k, p) \) be a random sample of size \( n \). Define \( Y = \sum_{i=1}^{n} X_i \), then \( M_Y(t) = \prod_{i=1}^{n} (q + pe^t)^k = (q + pe^t)^{kn} \).

**Example 2:** Let \( X_i \sim Gamma(1, \theta) \) be a random sample of size \( n \). Define \( Y = \sum_{i=1}^{n} X_i \), then \( M_Y(t) = \prod_{i=1}^{n} (1 - \theta t)^{-1} = 1/(1 - \theta t)^n \).

Exercises:
Random Functions Associated with Normal Distributions

In statistical applications, it is usually assumed that the population from which a sample is taken is $N(\mu, \sigma^2)$.

**Theorem:** Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from $N(\mu, \sigma^2)$. Define $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$, then $\bar{X} \sim N(\mu, \sigma^2/n)$.

**Theorem:** Let $X_j \sim \chi^2(r_j), 1 \leq j \leq n$. If $X_1, X_2, \ldots, X_n$ are independent, then $Y = \sum_{i=1}^{n} X_i \sim \chi^2(r_1 + r_2 + \ldots + r_n)$.

**Theorem:** Let $Z_1, Z_2, \ldots, Z_n$ be a random sample of size $n$ from $N(0, 1)$, then $W = Z_1^2 + Z_2^2 + \ldots + Z_n^2 \sim \chi^2(n)$.

**Corollary:** Let $\{X_i's\}$ be independent random variables from $N(\mu_i, \sigma_i^2)$, respectively, then $W = \sum_{i=1}^{n} (X_i - \mu_i)^2 / \sigma_i^2$ is $\chi^2(n)$.

**Theorem:** Let $\{X_i's\}$ be observations of a random sample of size $n$ from $N(\mu, \sigma^2)$. Define $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$, then

(a) $\bar{X}$ and $S^2$ are independent.

(b) $\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n} (X_i - \bar{X})^2 / \sigma^2 \sim \chi^2(n-1)$.

**Example 1:** Let $X_1, X_2, X_3, X_4$ be a random sample of size 4 from the normal distribution $N(76.4, 383)$. Then

(a) $U = \sum_{i=1}^{n} (X_i - 76.4)^2 / 383 \sim \chi^2(4)$, $P(0.711 \leq U \leq 7.779) = 0.90 - 0.05 = 0.85$.

(b) $W = \sum_{i=1}^{n} (X_i - \bar{X})^2 / 383 \sim \chi^2(3)$, $P(0.352 \leq W \leq 6.251) = 0.90 - 0.05 = 0.85$.

**Theorem:** Let $X_i \sim N(\mu_i, \sigma_i^2), 1 \leq i \leq n$, be independent. Define $Y = \sum_{i=1}^{n} a_i X_i$, then $Y \sim N(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2)$.

**Exercises:**
The Central Limit Theorem

**Theorem:** Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from $N(\mu, \sigma^2)$. Define $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$, then $\bar{X} \sim N(\mu, \sigma^2/n)$.

**Theorem:** Let $\bar{X}$ be the mean of a random sample $X_1, X_2, \ldots, X_n$ of size $n$ from a distribution with mean $\mu$ and variance $\sigma^2$. Define $W_n = (\bar{X} - \mu)/\left(\sigma/\sqrt{n}\right)$. Then

(a) $W_n = (\sum_{i=1}^{n} X_i - n\mu)/(\sqrt{n}\sigma)$

(b) $P(W_n \leq w) \approx \int_{-\infty}^{w} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \Phi(w)$.

(c) $W_n \sim N(0,1)$ as $n \to \infty$.

**Example 1:** Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from a $\chi^2(1)$. Define $Y = \sum_{i=1}^{n} X_i$. Then

(a) $Y \sim \chi^2(n)$.

(b) $(Y - n)/\sqrt{2n} \approx N(0,1)$.

**Example 2:** Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from a $U(0,1)$. Define $Y = \sum_{i=1}^{n} X_i$. Then

$(Y - 0.5n)/\sqrt{n/12} \approx N(0,1)$.

**Example 3:** Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from a $Bernoulli(p)$. Define $Y = \sum_{i=1}^{n} X_i$. Then

(a) $Y \sim b(n, p)$.

(b) $(Y - np)/\sqrt{np(1-p)} \approx N(0,1)$.

**Example 4:** Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from an exponential distribution with mean $\theta$. Define $Y = \sum_{i=1}^{n} X_i$. Then

(a) $Y \sim Gamma(n, \theta)$.

(b) $(Y - n\theta)/\sqrt{n\theta^2} \approx N(0,1)$.
Approximations for Discrete Distributions

Use the normal distribution to approximate probabilities for certain discrete-type distributions.

Example 1: Let $Y \sim b(10, 1/2)$. Then

$$
P(3 \leq Y < 6) = P(2.5 \leq Y \leq 5.5) = \Phi(0.316) - \Phi(-1.581) = 0.6240 - 0.0570 = 0.5670 \text{ (by Table II)}.
$$

Example 2: Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from a $Poisson(\lambda)$. Define $Y = \sum_{i=1}^{n} X_i$. Then

$$
(Y - n\lambda)/\sqrt{n\lambda} \approx N(0,1).
$$

Example 3: Let $Y \sim Poisson(\lambda = 20)$. Then

$$
P(16 < Y \leq 21) = P(16.5 \leq Y \leq 21.5) = P\left(\frac{16.5 - 20}{\sqrt{20}} \leq \frac{Y - 20}{\sqrt{20}} \leq \frac{21.5 - 20}{\sqrt{20}}\right) = \Phi(0.335) - \Phi(-0.783) = 0.4142
$$

Exercises:
Limiting Moment-Generating Functions

**Theorem:** If a sequence of moment-generating functions approaches a certain one, say, $M(t)$, then the limit of the corresponding distribution must be the distribution corresponding to $M(t)$.

**Example 1:** Let $Y \sim b(50, 0.04)$ and let $\lambda = np = 50 \times 0.04 = 2$. Then

\[
P(Y \leq 1) = 0.400
\]

\[
P(Y \leq 1) \approx 0.406 \text{ by a Poisson approximation.}
\]

**Exercises:**
Box-Muller Transformation

Let \( \{X_1, X_2\} \) be a random sample from \( U(0,1) \), define

\[
Z_1 = \sqrt{-2\ln X_1 \cos(2\pi X_2)}, \quad \text{and} \quad Z_2 = \sqrt{-2\ln X_1 \sin(2\pi X_2)},
\]

then

- The joint p.d.f. of \( X_1 \) and \( X_2 \) is \( f(x_1, x_2) = 1, \quad 0 < x_1, x_2 < 1 \),
- The joint p.d.f. of \( Z_1 \) and \( Z_2 \) is \( g(z_1, z_2) = \frac{1}{2\pi} \exp[-(z_1^2 + z_2^2)/2], \quad -\infty < z_1, z_2 < \infty \).
The Beta, Student’s t, and F Distributions

Random variables whose space are intervals or a union of intervals are said to be of the continuous types. The p.d.f. of a r.v. X of continuous type is an integrable function \( f(x) \) satisfying

(a) \( f(x) > 0, \ x \in R \)

(b) \( \int_{\mathbb{R}} f(x)dx = 1 \)

(c) The probability of the event \( X \in A \) is \( P(A) = \int_A f(x)dx \)

**Beta** \( f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}, \ 0 < x < 1; \ \alpha, \beta > 0 \)

**Student’s t** Let \( Z \sim N(0,1) \) and \( V \sim \chi^2(r) \) be two independent random variables. Define \( T = Z/\sqrt{V/r} \). Then \( T \) has a t-distribution with p.d.f.

\[
f(t) = \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\Gamma(r/2)} \frac{1}{(1+t^2/r)^{(r+1)/2}}, \quad -\infty < t < \infty
\]

**F-distribution** Let \( U \sim \chi^2(r_1) \) and \( V \sim \chi^2(r_2) \) be two independent random variables. Define \( W = (U/r_1)/(V/r_2) \). Then \( W \) has an F-distribution with p.d.f.

\[
f(w) = \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\Gamma(r_2/2)} \frac{w^{(r_1/2)-1}}{(1+w/(r_2/r_1))^{(r_1+r_2)/2}}, \quad 0 < w < \infty
\]

**Exercises:**
Expectation and Covariance Matrix

Let \( X_1, X_2, \ldots, X_n \) be random variables such that the expectation, variance, and covariance are defined as follows.

\[
\mu_j = E(X_j), \quad \sigma_j^2 = \text{Var}(X_j) = E[(X_j - \mu_j)^2]
\]

\[
\text{Cov}(X_i, X_j) = E((X_i - \mu_i)(X_j - \mu_j)) = \rho_{ij}\sigma_i\sigma_j
\]

Suppose that \( \mathbf{X} = [X_1, X_2, \ldots, X_n]^t \) is a random vector, then the expected mean vector and covariance matrix of \( \mathbf{X} \) is defined as

\[
E(\mathbf{X}) = [\mu_1, \mu_2, \ldots, \mu_n]^t = \mu, \quad \text{Cov}(\mathbf{X}) = [E((X_i - \mu_i)(X_j - \mu_j))]
\]
Multivariate (Normal) Distributions

◊ (Gaussian) Normal Distribution: $X \sim N(u, \sigma^2)$

\[
f_X(x) = f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-u)^2}{2\sigma^2}\right) \text{ for } -\infty < x < \infty
\]

mean and variance: $E(X) = u, \ Var(X) = \sigma^2$

◊ (Gaussian) Normal Distribution: $X \sim N(u, C)$

\[
f_X(x) = f(x) = \frac{1}{(2\pi)^{d/2}|\text{det}(C)|^{1/2}} \exp\left(-\frac{(x-u)^tC^{-1}(x-u)}{2}\right) \text{ for } x \in \mathbb{R}^d
\]

mean vector and covariance matrix: $E(X) = u, \ Cov(X) = C$

◊ Simulate $X \sim N(u, C)$

1. \( C = LL^t \), where $L$ is lower-$\Delta$.
2. Generate $y \sim N(0, I)$.
3. $x = u + L \ast y$
4. Repeat Steps (2) and (3) $M$ times.

% Simulate $N([1 \ 3]', [4,2; 2,5])$
% 
% n=30;
% X1=random('normal',0,1,n,1);
% X2=random('normal',0,1,n,1);
% Y=[ones(n,1), 3*ones(n,1)]+[X1,X2]*[2 1; 0, 2];
% Yhat=mean(Y) % estimated mean vector
% Chat=cov(Y) % estimated covariance matrix
% Z=[X1, X2];
Plot a 2D standard Gaussian Distribution

\begin{verbatim}
x=-3.6:0.3:3.6;
y=x';
X=ones(length(y),1)*x;
Y=y*ones(1,length(x));
Z=exp(-(X.^2+Y.^2)/2+eps)/(2*pi);
mesh(Z);
title('f(x,y)= (1/2\pi)*exp[-(x^2+y^2)/2.0]')
\end{verbatim}