Ch1. Review of Basic Probability and
Statistics Terminology

Outline of the Chapter:

- Some Basic Statistical Concepts
- Some Useful Discrete Distributions
- Some Useful Continuous Distributions

This chapter is a very brief review of basic properties and terminology from probability and statistics that we will use in this semester. The student is assumed to have seen most of this chapter before. This material is treated in most introductory texts on probability and statistics.
1 Some Basic Statistical Concepts

Given a real-valued random variable $X$, the Cumulative Distribution Function (c.d.f.) of $X$ is the function $F : \mathbb{R} \to [0, 1]$ defined by

$$F(x) = \Pr(X \leq x), \quad x \in \mathbb{R}.$$ 

If there exists a function $f : \mathbb{R} \to [0, \infty)$ such that $F(x) = \int_{-\infty}^x f(y) \, dy$ for every $x \in \mathbb{R}$, then $x$ is said to be continuous with Probability Density Function (p.d.f.) $f$. On the other hand, if $X$ only takes values in the set of integers, or more generally in some countable (or finite) set $S$, then its c.d.f. is completely determined by its Probability Mass Function (p.m.f.), $p : S \to [0, 1]$ where

$$p_i = \Pr(X = i), \quad i \in S.$$ 

Clearly, $\int_{-\infty}^{\infty} f(x) \, dx = 1$ for any p.d.f. and $\sum_{i \in S} p_i = 1$ for any p.m.f.
The Mean or Expectation of a real-valued random variable $X$ is defined by

$$E(X) = \begin{cases} 
\int_{-\infty}^{\infty} xf(x) \, dx & \text{if } X \text{ has p.d.f. } f \\
\sum_{i \in S} ip_i & \text{if } X \text{ has p.m.f. } p.
\end{cases}$$

The Variance of $X$, denoted by $\text{var}(X)$, is defined to be $E[(X - E(X))^2]$ and equals $E(X^2) - [E(X)]^2$ when it is finite. If we view $X$ as the value of some measurement, then the standard deviation $\sqrt{\text{var}(X)}$ determines the magnitude of error in this measurement.
Let $X_1, X_2, \cdots$ be a sequence of Independent Identically Distributed (iid) random variables with mean $\mu$ and variance $\sigma^2$. We define the $n$-th Sample Mean by

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Then $E(\bar{X}_n) = \mu$ and $\text{var}(\bar{X}_n) = \frac{\sigma^2}{n}$. When $n$ is sufficiently large, the Central Limit Theorem implies $\frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}}$ has an approximate $N(0, 1)$ distribution. Therefore,

$$Pr(\bar{X}_n - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}_n + 1.96 \frac{\sigma}{\sqrt{n}}) \approx 0.95$$

for sufficiently large $n$. When $\sigma^2$ is unknown, its value can be estimated by the Sample Variance

$$S^2_n = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2.$$

The sample variance is an Unbiased Estimator of $\sigma^2$ whenever the $X_i$'s are iid.
Consider two random variables $X$ and $Y$ (have some Joint Distribution). The Covariance of $X$ and $Y$ is defined by

$$\text{cov}(x, y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y).$$

If $X$ and $Y$ are Independent then $\text{cov}(X, Y) = 0$ (but the converse of this statement is false in general). It is easy to check that

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y)$$

The Correlation of $X$ and $Y$ is defined by the following normalization of covariance:

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}.$$
2 Some Useful Discrete Distributions

1. Bernoulli Let $p$ be a number in the interval $[0,1]$. The Bernoulli distribution with parameter $p$ is the discrete distribution on $\{0,1\}$ whose pmf is

$$p_0 = 1 - p, \quad p_1 = p.$$  

we write $X \sim \text{Bernoulli}(p)$ to say that the random variable $X$ has this distribution.

2. Binomial Let $n$ be a positive integer and let $p$ be a number in $[0,1]$. The binomial distribution with parameters $n$ and $p$ is the discrete probability distribution whose pmf is given by

$$p_i = \binom{n}{i} p^i (1 - p)^{n-i}, \quad i = 0, 1, \ldots, n.$$  

we write $X \sim B(n,p)$ to say that the random variable $X$ has this distribution. If $X_1, \ldots, X_n$ are iid with common distribution Bernoulli$(p)$, then $X_1 + \cdots + X_n \sim B(n,p)$. The $B(n,p)$ has mean $np$ and variance $np(1 - p)$. 
3. Geometric Let $p$ be a number in $[0,1]$. The geometric distribution with parameter $p$ is the discrete probability distribution whose pmf is given by

$$p_i = (1 - p)^{i-1}p, \ i = 1, 2, 3, \ldots$$

This describe the distribution of the number of tosses of a (possibly unfair) coin needed until the first “heads” appears. we write $X \sim \text{Geometric}(p)$ to say that the random variable $X$ has this distribution. The distribution has mean $1/p$ and variance $(1 - p)/p^2$.

4. Poisson Let $\lambda > 0$. The Poisson distribution with parameter $\lambda$ is the discrete probability distribution whose pmf is given by

$$p_i = \frac{e^{-\lambda} \lambda^i}{i!}, \ i = 0, 1, 2, \ldots$$

we write $X \sim \text{Poisson}(\lambda)$ to say that the random variable $X$ has this distribution. The mean and variance of this distribution are both equal to $\lambda$. 


3 Some Useful Continuous Distributions

**Lemma 3.1** Suppose that $X$ is a continuous random variable having pdf $f(x)$.

(i) If $a$ is a real number, then the pdf of $X + a$ is $f(x - a)$.

(ii) If $b$ is a positive number, then the pdf of $bX$ is $b^{-1}f(x/b)$.

1. **Uniform** Let $a$ and $b$ be real numbers with $a < b$. The uniform distribution on the interval $[a, b]$ is the continuous distribution whose pdf is given by

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

We write $X \sim U[a, b]$ to say that the random variable $X$ has this distribution.
2. **Normal** Let $\mu$ be a real number and $\sigma$ be a positive number. The normal distribution with mean $\mu$ and variance $\sigma^2$ is the continuous distribution whose pdf $f$ is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-(x - \mu)^2/2\sigma^2\right\}$$

we write $X \sim N(\mu, \sigma^2)$ to say that the random variable $X$ has this distribution. If $Z \sim N(0, 1)$, then $\mu + \sigma Z \sim N(\mu, \sigma^2)$.

3. **Exponential** Let $\lambda > 0$. The exponential distribution with parameter $\lambda$ is the continuous distribution whose pdf is given by

$$f(x) = \begin{cases} 
\lambda e^{-\lambda x} & \text{if } x \geq 0 \\
0 & \text{if } x < 0 
\end{cases}$$

we write $X \sim Exp(\lambda)$ to say that the random variable $X$ has this distribution. The distribution has mean $1/\lambda$ and variance $1/\lambda^2$. 
4. **Gamma** Let $a$ and $b$ be positive numbers. The gamma distribution with parameters $a$ and $b$ is the continuous distribution whose pdf $f$ is given by

$$f(x) = \begin{cases} \frac{b^a x^{a-1} e^{-bx}}{\Gamma(a)} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

where $\Gamma$ is the Gamma function

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} \, dt.$$  

The Gamma function satisfies the recursion $\Gamma(a) = (a-1)\Gamma(a-1)$. We write $X \sim \text{Gamma}(a, b)$ to say that the random variable $X$ has this distribution. This distribution has mean $a/b$ and variance $a/b^2$.

**Proposition 3.1** (i) Let $X_1, \ldots, X_k$ be independent random variables, and assume $X_i \sim \text{Gamma}(a_i, b)$ for each $i$. Then

$$X_1 + \cdots + X_k \sim \text{Gamma}(a_1 + \cdots + a_k, b).$$

(ii) Assume $Z \sim N(0, 1)$, then $Z^2 \sim \text{Gamma}(\frac{1}{2}, \frac{1}{2})$.

As a consequence of the above proposition, assume that $Z_1, \ldots, Z_k$ are iid $N(0, 1)$ random variables. Then $Z_1^2 + \cdots + Z_k^2$ has the $\text{Gamma}(\frac{k}{2}, \frac{1}{2})$ distribution. This arises often in statistics, and is called the chi-squared distribution with $k$ degrees of freedom, or $\chi^2_k$ for short.