STAT 302 Review of Chapters 5 – 7

(*November, 2012*)

1. Multivariate Probability Distributions

(a) Bivariate discrete distributions

- Joint distribution $p(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$.
- Marginal distribution $P(X_1 = x_1) = \sum_{x_2} p(x_1, x_2)$, $\quad P(X_2 = x_2) = \sum_{x_1} p(x_1, x_2)$.
- Conditional distribution $P(X_1 = x_1 | X_2 = x_2) = P(X_1 = x_1, X_2 = x_2)/P(X_2 = x_2)$.

Note: (i) $\sum_{x_1} \sum_{x_2} p(x_1, x_2) = 1$; (ii) $\sum_{x_1} P(X_1 = x_1 | X_2 = x_2) = 1$ for any fixed x_2. These results may be used to verify your answers.

(b) Bivariate continuous distributions

- Joint pdf $f(x_1, x_2)$: for any area A in the plane,

$$p((X_1, X_2) \in A) = \int \int_A f(x_1, x_2) dx_1 dx_2.$$

  (Please review techniques for double integrations.)

- Marginal pdf: $f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2$, $\quad f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_1$.
- Conditional pdf: $f(x_1 | x_2) = f(x_1, x_2)/f_2(x_2)$ for $f_2(x_2) > 0$.

Note: (i) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 = 1$. (ii) $\int_{-\infty}^{\infty} f(x_1 | x_2) dx_1 = 1$ for any fixed x_2. (These two results may be used to verify your answers.) (iii) $P(a < X < b \mid x_2) = \int_a^b f(x_1 | x_2) dx_1$ for any fixed x_2.

2. Expectation, Variance, Covariance, Correlation, and Independence.

a. Covariance and Correlation

Covariance and correlation describe the (linear) relationship between two random variables.

Covariance: $Cov(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2)$.

Correlation: $\rho(X_1, X_2) = \frac{Cov(X_1, X_2)}{\sqrt{V(X_1)V(X_2)}}$, $\qquad -1 \le \rho \le 1$.

Properties:

$$\begin{aligned}
E(h(X_1, X_2)) &= \int \int h(x_1, x_2) f(x_1, x_2) dx_1 dx_2, && \text{for any function } h(x_1, x_2), \\
V(\sum_{i=1}^n a_i X_i) &= \sum_{i=1}^n a_i^2 V(X_i) + 2 \sum \sum_{i<j} a_i a_j Cov(X_i, X_j), && \text{for any constants } a_i, \\
Cov(\sum_{i=1}^n a_i X_i, \sum_{j=1}^m b_j Y_j) &= \sum_{i=1}^n \sum_{j=1}^m a_i b_j Cov(X_i, Y_j), && \text{for any constants } a_i, b_j.
\end{aligned}$$

b. Independence

Random variables X_1 and X_2 are independent if $f(x_1, x_2) = f(x_1)f(x_2)$, for all x_1, x_2, when X_1 and X_2 are continuous, or $P(X_1 = x_1, X_2 = x_2) = P(X_1 = x_1)P(X_2 = x_2)$, for every x_1, x_2, when X_1 and X_2 are discrete. If X_1 and X_2 are independent, the $Cov(X_1, X_2) = 0$.

c. Functions of Random Variables

Let $Y = h(X)$, where $h(x)$ is a known function. We want to find the distribution of Y if the distribution of X is given.

- General approach: (i) If X is discrete,

$$P(Y = y) = \sum_{\{x:\, h(x)=y\}} P(X = x).$$

(ii) If X is continuous, first find the cdf of Y

$$F_Y(y) = P(Y \le y) = \int_{\{x:\, h(x)\le y\}} f_X(x)dx,$$

then find the pdf of Y via

$$f_Y(y) = dF_Y(y)/dy.$$

The above methods work for general problems, including bivariate case, but sometimes the methods can be tedious. In some cases, the following method may be preferable.

- If $h(x)$ is monotone and differentiable, then

$$f_Y(y) = f_X(h^{-1}(y))\Big|\frac{dx}{dy}\Big|.$$

- For multivariate cases, sometimes we can use the condition method illustrated in class.

d. Common distributions in statistics

The following distributions are widely used in statistics: the t-distribution, the χ^2-distribution, and the F distribution. It is important to know how these distributions are obtained, but it is not necessary to remember their density functions.

3. Limit Theorems

Sequences of random variables may converge in different ways: converge in probability, converge in distribution, etc.

X_n converges to X in probability if for any $\epsilon > 0$

$$P(|X_n - X| > \epsilon) \to 0.$$

X_n converges to X in distribution if $F_n(x)$ (the cdf of X_n) converges to $F(x)$ (the cdf of X) at every x where $F(x)$ is continuous. Remark: X_n converges to X in distribution if the mgf of X_n $M_n(t)$ converges to the mgf of X $M(t)$ at every t.

The Central Limit Theorem: If $X_1, \cdots, X_n$ are i.i.d. with $E(X_i) = \mu$ and $V(X_i) = \sigma^2 < \infty$, then

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \to N(0,1) \text{ in distribution.}$$

Note: This review may NOT include all the materials covered in class. The final exam will cover all the materials given in class.