Appendix B

Fundamentals of Probability

Joint Distributions, Conditional Distributions, and Independence

Joint Distributions

- Discrete Case: $f_{X,Y}(x,y) = P(X = x, Y = y)$
- Continuous Case: $f_{X,Y}(x,y)$ is its own function.

For *independent* random variables the joint distribution can be separated as the product of the marginal distributions.

- Discrete Case: $f_{X,Y}(x,y) = P(X = x) \cdot P(Y = y)$
- Continuous Case: $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$

Conditional Distributions

- Discrete Case: $f_{Y|X}(y|x) = P(X = x, Y = y)/P_x(X = x)$
- Continuous Case: $f_{Y|X}(y|x) = f_{X,Y}(x,y)/f_X(x)$

For *independent* random variables, knowing information about X does not change the probability of Y.

- Discrete Case: $P_{Y|X}(y|x) = P(Y = y) \cdot P(Y = y)/P(X = x) = P(Y = y)$
- Continuous Case: $f_{Y|X}(y|x) = f_Y(y) \cdot f_X(x)/f_X(x)$

Estimated Joint Probabilities of Discrete Variables: Crosstabulations

Sample proportions, or relative frequencies, estimate the population proportions (probabilities) of an outcome.

Re-create the crosstabulation of marital happiness and whether a couple has kids from the affairs dataset showing only the *proportion* (not percentage) of each joint outcome.

```
prop.table(table(affairs$marriage,affairs$haskids))
```

```
##
##                          no         yes
##   very unhappy 0.004991681 0.021630616
##   unhappy      0.013311148 0.096505824
##   average      0.039933444 0.114808652
##   happy        0.066555740 0.256239601
##   very happy   0.159733777 0.226289517
```

Estimated **Conditional Probabilities of Discrete Variables: Row Margins**

Conditional probabilities only consider the probablities of Y *given* a specific known outcome on X (or vice-versa). The `margins` option in `prop.table` conditions the relative frequencies on knowing the outcome of the row (`margins = 1`) or column (`margins = 2`).

Re-create the crosstabulation of marital happiness and whether a couple has kids from the affairs dataset showing only the *proportion* (not percentage) of each *conditional* outcome.

```
prop.table(table(affairs$marriage,affairs$haskids), margin = 1)
```

```
##
##                       no       yes
##   very unhappy 0.1875000 0.8125000
##   unhappy      0.1212121 0.8787879
##   average      0.2580645 0.7419355
##   happy        0.2061856 0.7938144
##   very happy   0.4137931 0.5862069
```

```
prop.table(table(affairs$marriage,affairs$haskids), margin = 2)
```

```
##
##                        no        yes
##   very unhappy 0.01754386 0.03023256
##   unhappy      0.04678363 0.13488372
##   average      0.14035088 0.16046512
##   happy        0.23391813 0.35813953
##   very happy   0.56140351 0.31627907
```

Covariance and Correlation

Covariance and correlation are (related) measures of the linear association between two variables.

Covariance

$$Cov(X,Y) = \sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)]$$

Distributing within the expected value, covariance "simplifies" to:

$$\sigma_{XY} = E(XY) - E(X)E(Y)$$

Variance is actually a special case of covariance, where $X = Y$.

$$V(X) = Cov(X,X) = \sigma_{XX} = E[(X-\mu_X)(X-\mu_X)]$$

Properties of Covariance

1. Adding a constant to either variable changes nothing: for any constants, a_1 and a_2, $Cov(a_1+X, a_2+Y) = Cov(X,Y)$
2. Multiplication factors out: for any constants, a_1, a_2, b_1, and b_2, $Cov(a_1+b_1X, a_2+b_2Y) = b_1b_2Cov(X,Y)$
3. Linear functions: for any constants, b_1 and b_2, $Cov(b_1X, Y) = b_1Cov(X,Y)$, $Cov(X, b_2Y) = b_2Cov(X,Y)$, and $Cov(b_1X, b_2Y) = b_1b_2Cov(X,Y)$
4. If X and Y are independent, $Cov(X,Y) = 0$).
5. Covariance is bounded: $|Cov(X,Y)| \leq \sigma_X \cdot \sigma_Y$

Covariance Example

Using the ceosal1 data, calculate:

- $E(Salary \cdot ROE)$ (name this "E_sal.roe")
- $E(Salary) \cdot E(ROE)$ (name this "E_sal.E_roe")
- $Cov(Salary, ROE) = E(Salary \cdot ROE) - E(Salary) \cdot E(ROE)$

```
attach(ceosal1)
E_sal.roe <- mean(salary*roe)
E_sal.E_roe <- mean(salary)*mean(roe)
E_sal.roe - E_sal.E_roe
```

```
## [1] 1336.115
```

Correlation Coefficient

The correlation coefficient is a standardized measure of association.

$$Corr(X,Y) = \rho_{XY} = \frac{Cov(X,Y)}{\sqrt{V(X) \cdot V(Y)}} = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y}$$

Correlation Example

```
(E_sal.roe - E_sal.E_roe)/(sd(salary)*sd(roe))
```

```
## [1] 0.1142923
```

Using `cov` and `cor` Functions

Calculate the covariance and correlation for salary and roe using the "cov" and "cor" functions.

```
cov(ceosal1$salary, ceosal1$roe)
```

```
## [1] 1342.538
```

```
cor(ceosal1$salary, ceosal1$roe)
```

```
## [1] 0.1148417
```

Correlation and Independence

- What is $corr(x, x^2)$?
- Are x and x^2 independent?",

Conditional Expectation

For a discrete random variable, $Y \in \{y_1, y_2, ..., y_m\}$, the conditional expected value of Y is given by

$$E(Y|X) = \sum_{j=1}^{m} y_j f_{Y|X}(y_j|x)$$

For a continuous random variable,

$$E(Y|X) = \int_{y \in \Omega_Y} y_j f_{Y|X}(y|x) dy$$

Properties of Conditional Expectation

1. For any function, $c(X)$, $E[c(X)|X] = c(X)$
2. For *any* functions, $a(X)$ and $b(X)$, $E[a(X)Y + b(X)|X] = a(X)E(Y|X) + b(X)$ For example, $E(XY + 2x^2|X) = XE(Y|X) + 2X^2$
3. If X and Y are independent, $E(Y|X) = E(Y)$
4. If $E(Y|X) = E(Y)$, then $Cov(X,Y) = 0$ *and* for *any* function, $f(X)$, $Cov(f(X), Y) = 0$
5. If $E(Y^2) < \infty$ and for some function g, $E(g(X)^2) < \infty$ then:

a. The conditional mean is better (based on expected squared prediciton error) than any other function of X for predicting Y, conditional on X: $E[(Y - \mu_X)^2|X] \leq E\{[Y - g(X)]^2|X\}$ for any function $g(X) \neq \mu_X$
b. The conditional mean minimizes the unconditional expected squared prediction error: $E[(Y - \mu_X)^2] \leq E\{[Y - g(X)]^2\}$ for any function $g(X) \neq \mu_X$

6. Law of Iterated Expectations

a. The expected value over X of the expected value of Y *conditional on* X over all values of X equals the unconditional mean of Y: $E[E(Y|X)] = E(Y)$.
b. Given random variables X and Z, we can find the expected value of Y conditional on X in two steps. i. Find $E(Y|X, Z)$ ii. Take the expected value of the result conditional on X

$$E(Y|X) = E[E(Y|X, Z)|X]$$

Conditional Variance

$$V(Y|X = x) = E(Y^2|x) - [E(Y|x)]^2$$

Normal (and Other) Distributions

Univariate Normal Distribution

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{\frac{-(x-\mu_x)^2}{2\sigma_x^2}}$$

Standard Normal Distribution

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$$

Bivariate Normal Distribution (Independent Ranodm Variables):

$$f(x,y) = f(x) \cdot f(y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}([\frac{(x-\mu_x)}{\sigma_x}]^2 + [\frac{(y-\mu_y)}{\sigma_y}]^2)}$$

Bivariate Normal Distribution (Correlated Random Variables):

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{\frac{-1}{2(1-\rho^2)}[(\frac{x-\mu_x}{\sigma_x})^2 - 2\rho[(\frac{x-\mu_x}{\sigma_x})(\frac{y-\mu_y}{\sigma_y})] + (\frac{y-\mu_y}{\sigma_y})]^2}$$

Multivariate Normal Distribution (Matrix Form):

$$f(\mathbf{x}) = \frac{1}{\sqrt{2\pi|\mathbf{\Sigma}|}} e^{(\mathbf{x}-\mu)'\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)}$$

Other Distributions

- Student's t Distribution
- Chi-Square Distribution
- F Distribution