Review of Probability and Statistics: Key Terms

Basic Terminology

1. **Random Variables:** X, Y, Z, … vs. observations *x, y, z*, …

2. **Expectation**: $E(X) = \mu = \sum x\ f_X(x)$; $E(aX + b) = aE(X) + b$;

3. **Variance**: $Var(X) = \sigma^2 = E(X-\mu)^2 = \sum (x-\mu)^2 f_X(x)$; $Var(aX + b) = a^2 Var(X)$

4. **Standard Deviation**: $\sigma = \sqrt{\sigma^2}$, $StdDev(aX + b) = \sqrt{a^2 Var(X)} = a\ StdDev(X)$

5. **Standardize a Random Variable**: $Z = \dfrac{X-\mu}{\sigma}$; $E(Z) = 0$ and $Var(Z) = 1$

6. **Joint Density**: $f_{XY}(x, y) = P(X = x \,\&\, Y = y)$

7. **Independence**: $f_{XY}(x, y) = P(X = x)P(Y = y) = f_X(x) f_Y(y)$ for all *x* and *y*

8. **Covariance**: $Cov(X,Y) = \sigma_{XY} = E(X-\mu_X)(Y-\mu_Y) = \sum (x-\mu_X)(y-\mu_Y) f_{XY}(x, y)$;

   $Cov(X, X) = \sigma_{XX} = \sigma_X^2 = Var(X)$, $Cov(a + bX, c + dY) = bdCov(X,Y)$

9. If X and Y are independent, then $Cov(X,Y) = \sigma_{XY} = 0$; opposite need not hold

10. **Correlation**: $Corr(X,Y) = \rho_{XY} = \dfrac{Cov(X,Y)}{StdDev(X)\,StdDev(Y)} = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$, $-1 \le \rho_{XY} \le 1$

11. **Conditional Density**: $f_{Y|X}(y \mid x) = P(Y = y \mid X = x) = \dfrac{f_{X,Y}(x, y)}{f_X(x)}$

12. **Law of Iterated Expectations I**: $E\left[g(X,Y)\right] = E_X\left\{E_{Y|x}\left[g(x,Y) \mid x\right]\right\}$

13. If X and Y are independent then $f_{Y|X}(y \mid x) = f_Y(y)$ and $f_{X|Y}(x \mid y) = f_X(x)$

Multiple RV's: X, Y, and $\{X_i\}$

14. **Expectation**: $E(a_0 + \sum a_i X_i) = a_0 + \sum E(a_i X_i) = a_0 + \sum a_i E(X_i)$

15. **Variance**: $Var(X+Y) = \sigma_X^2 + 2Cov(X,Y) + \sigma_Y^2$, $Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j)$

16. **Independence**: $f_{X_1,\ldots,X_n}(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_n}(x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$

17. If $Cov(X_i, X_j) = 0$ when $i \neq j$, then $Var\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j) = \sum_{i=1}^{n} a_i^2 \sigma_i^2$

18. $Corr(a_1 X_1 + b_1, a_2 X_2 + b_2) = \pm Corr(X_1, X_2)$, $sign = sign(a_1 a_2)$

19. **Law of Iterated Expectations II**:

$$E\left[g(X_1, \ldots X_n, Y)\right] = E_{X_1,\ldots X_n}\left\{E_{Y|x_1\ldots x_n}\left[g(x_1, \ldots x_n, Y) \mid x_1, \ldots x_n\right]\right\}$$

Normal Distribution

20. **Standard Normal** (Gaussian): $\mathrm{N}\left(\mu, \sigma^2\right)$ has mean μ and variance σ^2

21. **Independent and Identically Distributed:** *iid*

22. If $\{X_i\}$ *iid* $N(\mu, \sigma^2)$ and $Y = \sum \alpha_i X_i$, then Y will have a Normal Distribution, with

$$E(Y) = \sum \alpha_i \mu = \mu \sum \alpha_i \text{ and } Var(Y) = \sum \alpha_i^2 \sigma^2 = \sigma^2 \sum \alpha_i^2 .$$

23. If $\{X_i\}$ *iid* $N(\mu, \sigma^2)$ and $\bar{X} = \frac{1}{n}\sum X_i$, then $\bar{X} \sim \mathrm{N}(\mu, \frac{1}{n}\sigma^2)$ and $Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$