STAT 302. Review of Chapters 1 to 4

1. Set notation for random events

Let A, B, C be random events. Then

$A \cup B$: either A or B or both occur (at least one event occurs)

AB: Both A and B occur (A and B occur together)

$\bar{A}$: event A does not occur

DeMorgan's Laws: $\overline{A \cup B} = \bar{A}\bar{B}$, $\overline{AB} = \bar{A} \cup \bar{B}$.

2. Counting Rules

A simple formula: if the total number of outcomes is finite and each outcome is equally likely, then

$$P(A) = \frac{\text{number of outcomes in A}}{\text{total number of outcomes}}.$$

The following counting rules will be useful for counting the number of outcomes in the above formula.

(a) **Multiplication Principle**: if we can breakdown a task into k steps, with step i having n_i outcomes, $i = 1, \cdots, k$, then there are a total of $n_1 n_2 \cdots n_k$ outcomes.

(b) **Permutation Rule and Combination Rule**: When selecting r objects from n objects, (i) the number of ordered arrangements is

$$P_r^n = \frac{n!}{(n-r)!}; \qquad \text{(Permutation Rule)}$$

(ii) the number of unordered arrangements is

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}. \qquad \text{(Combination Rule)}$$

3. Conditional Probability

$P(A|B) =$ probability of event A given that event B has already occurred.

$$P(A|B) = \frac{P(AB)}{P(B)}.$$

In general $P(A|B) \neq P(A)$. But if $P(A|B) = P(A)$, we call events A and B *independent.*

4. Probability Rules

Additive Rule: (i) General result: $P(A \cup B) = P(A) + P(B) - P(AB)$. (ii) If A and B are *mutually exclusive* (i.e., $AB = \phi$), we have $P(A \cup B) = P(A) + P(B)$.

Multiplicative Rule: (i) General result: $P(AB) = P(A)P(B|A) = P(B)P(A|B)$. (ii) If A and B are *independent*, we have $P(AB) = P(A)P(B)$.

Bayes Rule: If events $B_1, B_2, \cdots, B_k$ form a partition of the sample space, then for any j

$$P(B_j|A) = \frac{P(B_j)P(A|B_j)}{\sum_{i=1}^k P(B_i)P(A|B_i)}.$$

Note: the following formula is sometimes useful

$$P(A) = \sum_{i=1}^k P(B_i)P(A|B_i).$$

5. Discrete Random Variables

(i) A random variable is determined by all its possible values *and* by the probability associated with each value. We may use either the probability distribution, $P(X = x)$, for all x, or the cdf $F(x) = P(X \le x)$, for all x, to describe a random variable.

(ii) The expectation of a random variable is its "average" value, while the variance (or standard deviation) describes the variability (or how spread-out) of a random variable.

$$\begin{aligned} E(X) &= \sum_{i=1}^{\infty} x_i P(X = x_i), \\ V(X) &= E(X - E(X))^2 = \sum_{i=1}^{\infty}(x_i - E(X))^2 P(X = x_i) = E(X^2) - (E(X))^2. \\ \text{Standard deviation} &= \sqrt{V(X)}. \end{aligned}$$

Tchebysheff's theorem.

6. Common Discrete Distributions

Consider: in what situations does each of these distributions naturally arise, and in these situations what are the values of the parameters in the distributions (e.g., values of p, n, λ, etc).

(i) **Binomial Distribution:** Let X be the *number* of successes out of n independent Bernoulli trials and p be the probability of success in each Bernoulli trial. Then

$$\begin{aligned} P(X = k) &= \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \cdots, n. \\ E(X) &= np, \qquad V(X) = np(1-p). \end{aligned}$$

(ii) **Geometric Distribution:** Let X be the trial at which the *first* success occurs in a sequence of independent Bernoulli trials with p the probability of success in each Bernoulli trial. Then

$$\begin{aligned} P(X = k) &= (1-p)^{k-1} p, \qquad k = 1, 2, 3, \cdots \\ E(X) &= 1/p, \qquad V(X) = (1-p)/p^2. \end{aligned}$$

(iii) **Poisson Distribution:** Here X often represents the number of items (people, accidents, errors, etc) in a certain area or a certain time interval. Then (approximately)

$$\begin{aligned} P(X = k) &= \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k = 0, 1, 2, \cdots \\ E(X) &= \lambda, \qquad V(X) = \lambda. \end{aligned}$$

7. Continuous Random Variables

For a continuous random variable X, its probability distribution is determined by its pdf $f(x)$ which satisfies

$$P(a < X < b) = \int_a^b f(x)dx, \qquad \text{for any real numbers } a, b.$$

The cdf of X is given by

$$F(x) = \int_{-\infty}^x f(t)dt, \qquad -\infty < x < \infty.$$

Note: $F'(x) = f(x)$. This can be used to verify your answers.

Note: We always have (a) $f(x) \geq 0$, and $\int_{-\infty}^{\infty} f(x)dx = 1$. (b) $F(x)$ is non-decreasing, $0 \leq F(x) \leq 1$, $F(-\infty) = 0$, $F(+\infty) = 1$. These results may be used to verify your answers.

Expectation and Variance

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x)dx, \\ V(X) &= E(X - E(X))^2 = E(X^2) - (E(X))^2. \end{aligned}$$

Properties: $E(aX + b) = aE(X) + b, \quad V(aX + b) = a^2V(X)$.

8. Common Continuous Distributions

(a) Uniform distribution over $[a, b]$. Density (pdf) is given by

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

Note: $E(X) = (b + a)/2, \qquad V(X) = (b - a)^2/12$.

(b) The exponential distribution. Density (pdf) is given by

$$f(x) = \begin{cases} \frac{1}{\theta}e^{-\frac{1}{\theta}x}, & x \geq 0 \\ 0, & \text{elsewhere.} \end{cases}$$

Note: $E(X) = \theta, \qquad V(X) = \theta^2$.

(c) The Normal distribution. Density (pdf) is given by

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$

Note: $f(x)$ is symmetric about $x = \mu$. $E(X) = \mu, \qquad V(X) = \sigma^2$.

Remarks: (i) When $\mu = 0$, $\sigma^2 = 1$, we have the *standard normal distribution* $Z \sim N(0, 1)$. (ii) If $X \sim N(\mu, \sigma^2)$, then X can be standardized as follows

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

(iii) If $X \sim N(\mu, \sigma^2)$, then $aX + b \sim N(a\mu + b, a^2\sigma^2)$. (iv) To calculate normal probabilities $P(a < X < b)$, you need to standardize the variable X first, then use a probability table for $N(0, 1)$.