Summary for Finals

STAT 3013, Introduction to Probability

May 2, 2017

Basic Properties

1. For any event A of sample space Ω, the probability of A, denoted as $P(A)$ satisfies (1) $0 \leq P(A) \leq 1$, (2) $P(\Omega) = 1$ and (3) for mutually exclusive events, A_i, $i \geq 1$, $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

2. $P(A^c) = 1 - P(A)$. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. (Principle of Inclusion-Exclusion or PIE).

Conditional Probability and Independence

1. For any two events E and F, the conditional probability of E given F is denoted by

$$P(E \mid F) = P(E \cap F)/P(F).$$

2. $P(E) = P(E \mid F)P(F) + P(E \mid F^c)P(F^c)$.

3. **Bayes rule** Let F_i, $i = 1, \ldots, n$ be mutually exclusive events whose union is the sample space Ω. Then:

$$P(F_j \mid E) = \frac{P(E \mid F_j)P(F_j)}{\sum_{i=1}^{n} P(E \mid F_i)P(F_i)}$$

4. Two events E and F are said to be independent if they satisfy $P(E \cap F) = P(E)P(F)$.

Discrete Random Variable

1. A real-valued function defined on the outcome of a probability experiment is called a random variable.

2. For *any* random variable X, the distribution function is $F(x) = P(X \leq x)$.

3. A random variable whose set of possible values is either finite or countably infinite is called discrete. For a discrete random variable: $p(x) = P(X = x)$ is called the probability mass function of X. Expectation is defined as $E(X) = \sum_{x:p(x)>0} xp(x)$.

4. For any function g, $E(g(X)) = \sum_{x:p(x)>0} g(x)p(x)$.

5. The variance of *any* random variable X, denoted by $Var(X)$, is defined by:

$$Var(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

6. For k random variables $X_1, \ldots, X_k$ and constants $c_1, \ldots, c_k$,

$$E(\sum_{i=1}^{k} c_i X_i) = c_i E(X_i)$$

7. If in addition, $X_1, X_2, \ldots, X_k$ are independent:

$$Var(\sum_{i=1}^{k} c_i X_i) = c_i^2 Var(X_i)$$

8. **Tail sum method** For a nonnegative integer-valued random variable:

$$E(X) = \sum_{n=0}^{\infty} P(X > n)$$

Standard Discrete Distributions

1. **Binomial**(n, p) Number of heads in n coin tosses, each with probability p:

$$\text{PMF: } p(x) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, \ldots, n.$$

$$E[X] = np, Var[X] = np(1-p).$$

2. **Geometric(**(p) Number of tosses to get the first head, $P(head) = p$.

$$\text{PMF: } p(x) = p(1-p)^{x-1}, x = 1, 2, \ldots,$$
$$E[X] = 1/p, Var[X] = (1-p)/p^2.$$

3. Geometric is memoryless: $P(X > m+n \mid X > m) = P(X > n)$.

4. **Negative Binomial**(r, p) Number of tosses to get r heads, $P(head) = p$. Neg-Bin(r, p) is sum of r $Geom(p)$ r.v.s.

$$\text{PMF: } p(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, x \geq r$$
$$E[X] = r/p, Var[X] = r(1-p)/p^2.$$

5. **Poisson**(λ) Used to model number of events.

$$\text{PMF: } p(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x \geq 0$$
$$E[X] = \lambda, Var[X] = \lambda$$

6. **Hypergeometric**(n, N, m) Number of white balls selected when n balls are randomly chosen from an urn that contains N balls of which m are white. Let $p = \frac{m}{N}$.

$$\text{PMF: } p(x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}, x = 0, \ldots, m$$
$$E[X] = np, Var[X] = \frac{N-n}{N-1} np(1-p).$$

7. An important property of the expected value is that the expected value of a sum of random variables is equal to the sum of their expected values:
$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

8. If $X \equiv X_n \sim Bin(n, p), n \to \infty, p = p_n \to 0, np_n \to \lambda$ for some $0 < \lambda < \infty$, $X_n \approx$ Poisson(λ). This is called the **Poisson approximation to Binomial**. Useful for rare events modeling.

Continuous Distributions

1. A random variable X is continuous if there is a nonnegative function f , called the probability density function of X, such that, for any set B,
$$P(X \in B) = \int_B f(x)dx \equiv P(a \leq X \leq b) = \int_a^b f(x)dx$$

2. For a density f: $\int_{-\infty}^{\infty} f(x)dx = 1$. Useful for validating if a given function is a density and for finding normalizing constants.

3. Continuous random variable puts zero probability to discrete sets or single points.

4. If X is continuous, its CDF $F(x)$ is differentiable and
$$\frac{d}{dx}F(x) = f(x)$$

5. The expected value of a continuous random variable X is defined by $E[X] = \int_{-\infty}^{\infty} x f(x)dx$.

6. For any function g: $E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$.

7. **Uniform**(a, b): Denoted by $X \sim \mathcal{U}(a, b)$

$$\text{PDF: } p(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$
$$E[X] = \frac{b+a}{2}, Var[X] = \frac{(b-a)^2}{12}$$

8. **Normal**(μ, σ): Denoted by $X \sim \mathcal{N}(\mu, \sigma)$

$$\text{PDF: } p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$
$$E[X] = \mu, Var[X] = \sigma^2.$$

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

10. $P(X \leq x) = P(Z \leq \frac{x-\mu}{\sigma}) \doteq \Phi(\frac{x-\mu}{\sigma})$.

11. Probabilities about X can be expressed in terms of probabilities about the standard normal variable Z, obtained from **Normal Table**.

12. **Central Limit Thoerem:** When n is large, the probability distribution function of a binomial random variable with parameters n and p can be approximated by that of a normal random variable having mean np and variance $np(1p)$.