Basic Probability and Statistics: BMED 6517

Saurabh Sinha

saurabh.sinha@bme.gatech.edu
http://www.sinhalab.net

BME & ISyE
Georgia Tech

December 26, 2022

References

- Chapter 2.1-2.9 of Polanski & Kimmel.
- Chapter 11 of Durbin *et al.*

Some Basic Probability Rules

- If B and C are mutually exclusive and exhaustive, then $P(A) = P(A, B) + P(A, C)$

- $Pr(A) = \sum_{k=1}^{\infty} Pr(A \cap B_k)$, where the B_j's are exhaustive (union is the entire sample space) and mutually exclusive (no two of them overlap)

- If A and B are independent, then $P(A, B) = P(A)P(B)$

- More generally (regardless of independence of A, B), $P(A, B) = P(A)P(B|A) = P(B)P(A|B)$, where $P(X|Y)$ is the conditional probability of X given (conditional on) Y.

- **Point to remember**: Know when probabilities are added and when they are multiplied.

Conditional Probability

https://www.youtube.com/watch?v=H02B3aMNKzE

Random Variables

- Random variables are mappings from the sample space Ω to the space R of real numbers.
- Discrete random variable X takes values $x_0, x_1, \ldots x_k \ldots$ with corresponding probabilities $p_0, p_1, \ldots p_k \ldots$, with the condition

$$\sum_{k=0}^{\infty} p_k = 1$$

.

- The *distribution* of X.

Joint probability distributions

- $p_{ij} = P[X = x_i, Y = y_j]$
  with $\sum_{i,j} p_{ij} = 1$
- $p_i = P[X = x_i] = \sum_j p_{ij}$
  The *marginal* probability distribution of X.

Expectation and Moments

- Expectation of a function $g(x)$ with respect to the distribution of X is

$$E[g(X)] = \sum_{k=0}^{\infty} p_k g(x_k)$$

- When $g(x) = x$, we get the expectation of X: this is the "first moment" of X.

- Higher order moments: $E[g(x)]$ with $g(x) = x^n$.

- Variance: second order "central" moment: $E[(X - E(X))^2]$

- Linearity of expectation: $E(X + Y) = E(X) + E(Y)$.

- Not true for variance, except when X and Y are independent.

Discrete Distributions I

- Bernoulli trials and Binomial distribution.
- Bernoulli trial is an experiment with two possible outcomes: e.g., “success” with prob. p and “failure” with prob. $1-p$.
- Suppose you do K independent Bernoulli trials, each with same success prob. What is the distribution of the number of successes?
- Binomial distribution: prob. of k successes in K trials.

$$p_k = \binom{K}{k} p^k (1-p)^{K-k}$$

- Mean $E(X) = Kp$; Variance $\mathrm{Var}(X) = Kp(1-p)$
- Example: You sell sandwiches. 40% of customers order chicken, 60% order something else. What is the probability that at least four of the next five sandwiches you sell will be chicken? (Answer: 8.7%)

Discrete Distributions II

- Geometric distribution
- Probability that in a series of Bernoulli trials, the first success happens exactly after k failures.

$$p_k = (1-p)^k p$$

- $E(X) = \frac{1-p}{p}$ and $\text{Var}(X) = \frac{1-p}{p^2}$
- Example: 10% of customers order a veggie sandwich. How many sandwiches do you expect to sell before the next veggie order comes along? (Answer: 9)

Discrete Distributions III

- Poisson distribution
- Used to model the number of occurrences of events that happen at random moments, over a fixed interval of time, e.g., number of 911 calls.

$$p_k = \frac{e^{-\lambda}\lambda^k}{k!}$$

- λ is a parameter.
- $E(X) = \lambda$, $\mathrm{Var}(X) = \lambda$
- You get about two sandwich orders a minute. What is the chance that you'll get five or more orders in the next minute? (Answer: 5.3%)

Continuous Distributions

- Density function (pdf) $f(x)$ and cumulative prob. distribution $F(x) = P[X \leq x]$.

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

- $E[g(x)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$

- Higher moments are similarly defined.

- With continuous distributions, probability of any particular value is 0. We talk about "probability density" at a particular value, not its probability.

Continuous distribution I

- Normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \text{exp} \left[-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right]$$

- μ and σ are parameters equal to expectation and standard deviation resp.

- Supported on the whole set of reals

- The famous Bell curve!