Review of 120A

Pstat 120B

Probability Measure (Probability Function)

Given a sample space $\mathbf{S}$, a probability function $P(\cdot)$ is a function that satisfies the following three conditions:

- Non-negative: $P(A) \geq 0$ for any set $A \in \mathbf{S}$
- Total probability of 1: $P(\mathbf{S}) = 1$
- Countable Additivity: Let $A_1, A_2, \ldots$ be a pairwise mutually exclusive (disjoint) events in $\mathbf{S}$, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Note: Pairwise mutually exclusive events means that $A_i \cap A_j = \emptyset$ for $i \neq j$.

Empty set: $P(\emptyset) = 0$

Complements: $P(A^c) = 1 - P(A)$.

Subsets: If $A \subset B$, then

$$P(A) \leq P(B).$$

Valid Probabilities: Let $A \in \mathbf{S}$, then

$$0 \leq P(A) \leq 1.$$

Additive Rule: For events $A, B \in \mathbf{S}$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Conditional Probability: For events $A, B \in \mathbf{S}$, the conditional probability of A happening given that B has already occurred is given by

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

provided that $P(B) \neq 0$.

Multiplicative Rule: For events $A, B \in \mathbf{S}$

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

Independance: Events $A, B \in \mathbf{S}$ is said to be independent if

$$P(A \cap B) = P(A)P(B).$$

Bayes' Theorem: For events $A, B \in \mathbf{S}$

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}.$$

Random Variable (RV):

A random variable is a mapping (function) from the sample space, **S**, to the real numbers, $\mathbb{R}$. Random variables are usually denoted by capital letters, X, and the possible values that they can take on are denoted by lower case letters, x. You may think of a random variable as a convenient placeholder for various probability calculations.

Example 1:
Suppose you're interested in the number of accidents occurring on the 101 today

Let X = the number of accidents that occur on the 101

Now rather than writing statements like "the probability of 4 accidents occuring is 0.2" or "the expected number of accidents is 2.3" you may simply write $P(X = 4) = 0.2$ and $E(X) = 2.3$

Support:

The support of a random variable is defined to be the set of all possible values which the random variable can attain.

Discrete vs. Continuous Random Variables:

A discrete random variable is one with finite or countably infinite support (e.g. $\{-4, -2, 0, 2, 4\}$ or $\{0, 1, 1, 2, 3, 5, 8, 13, \ldots\}$) and a continuous random variable has uncountable support (e.g. $(0, 1)$ or $(0, \infty)$). So if the possible values that the random variable can take on is a list of numbers it's discrete and if it's an interval then it's continuous.

Discrete random variables can characterize their probabilities with a probability mass functions (PMF) whereas, continuous random variable use probability density functions (PDF). The probability mass function (PMF) of a discrete random variable, is defined by:

$$p_X(x) = P(X = x).$$

To check that a function is a legitimate probability function, you need to check that:

- $p_X(x) \geq 0$ for all x
- $\sum_{\text{all } x} p_X(x) = 1$

The probability density function (PDF) of a continuous random variable is given by $f_X(x)$, where:

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx.$$

To check that $f_X(x)$ is a legitimate PDF, you need to check that:

- $f_X(x) \geq 0$ for all x
- $\int_{-\infty}^{\infty} f_X(x) = 1$

Joint Distributions: Discrete case

Suppose that X and Y are two discrete random variables defined on a sample space, S. Then the joint probability function is given by:

$$p_{X,Y}(x, y) = P(X = x, Y = y).$$

Note: for a set of points, A, the probability of A happening is

$$P(A) = \sum_{\text{all } (x,y) \in A} \sum p_{X,Y}(x, y)$$

Discrete Marginal Distributions:

$$p_X(x) = P(X = x) = \sum_{\text{all } y} p_{X,Y}(x, y)$$
$$p_Y(y) = P(Y = y) = \sum_{\text{all } x} p_{X,Y}(x, y)$$

Joint Distributions: Continuous Case

Suppose that X and Y are two continuous random variables defined on a sample space, **S**. Then the joint probability density function (joint PDF) is given by $f_{X,Y}(x, y)$, where:

$$P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f_{X,Y}(x, y)dydx$$

Continuous Marginal Distributions:

$$f_X(x) = \int_{\text{all } y} f_{X,Y}(x, y)dy$$
$$f_Y(y) = \int_{\text{all } x} f_{X,Y}(x, y)dx$$

Independent Random Variables

Theorem: Two random variables X and Y are independent if and only if

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

Corollary: If $X_1, X_2, \ldots, X_n$ are independent random variables, then

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

Cumulative Distribution Function (CDF)

$$F_X(x) = P(X \leq x)$$

Discrete case: If you have a discrete random variable, then

$$F_X(x) = P(X \leq x) = \sum_{k=-\infty}^{x} p_X(k)$$

Continuous case: if you have a continuous random variable, then

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt$$

Relationship Between PDFs and CDFs

If X is a continuous random variable, then by the Fundamental Theorem of Calculus

$$\frac{d}{dx}F_X(x) = \frac{d}{dx}\int_{-\infty}^{x} f_X(t)dt = f_X(x)$$

Conditional Distribution:

Let X, Y be jointly distributed with joint PDF $f_{X,Y}(x,y)$. Then the conditional distribution of X given that $Y=y$ is given by

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

provided that $f_Y(y) > 0$.

Expected Value: ("average")

$$\text{Discrete case:} \qquad \mu = E[X] = \sum_{\text{all } x} x p_X(x)$$

$$\text{Continuous case:} \qquad \mu = E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

Expectation of a Function, $g(X)$

$$\text{Discrete case:} \qquad E[g(X)] = \sum_{\text{all } x} g(x) p_X(x)$$

$$\text{Continuous case:} \qquad E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

Moments

The kth moment of a random variable, X is defined as

$$E(X^k)$$

Theorem:

1. If X and Y are independent, then
$$E(XY) = E(X)E(Y).$$
2. If a and b are constants, then
$$E(aX+b) = aE(X) + b.$$
3. Let $X_1, X_2, \ldots, X_n$ be any random variables, then
$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$
4. Tower Property:
$$E[X] = E[E(X|Y)]$$

Variance ("spread")

The variance of a random variable is defined to be

$$Var(X) = E\left[(X - E[X])^2\right] = E\left[(X-\mu)^2\right]$$

Theorem:

1. $Var(X) = E(X^2) - (E[X])^2 = E[X^2] - \mu^2$
2. Let X and Y be any random variables. If a and b are constants, then

$$Var(aX + b) = a^2 Var(X)$$
$$Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y)$$

3. Let $X_1, X_2, \ldots, X_n$ be independent random variables, then

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

Moment Generating function:

The moment generating function (MGF) of a random variable, X is defined as

$$M_X(t) = E(e^{Xt})$$

Theorem:

1. If X and Y are independent, then

$$M_{X+Y}(t) = M_X(t)M_Y(t).$$

2. If c is a constant, then

$$M_{cX}(t) = M_X(ct).$$

3. The kth derivative of the MGF evaluated at 0 gives the kth moment of the random variable.