MA-224 Review

Jacob Sigman

Contents

1 Probability

1.1 Basic Concepts

Statistics The study of collected data.

Data Qualitative or quantitative information that you obtain through experiments.

Experiment An unbiased, random, planned activity. Outcomes are simple, events are more complex.

Frequency The number of times that something occurs. There are three ways to display frequency: A frequency table, a bar graph, and a relative frequency table.

Probability Experimental and theoretical are two types. Theoretical is if an experiment is done many times. It is equivalent to $\frac{\text{\# of times event occurs}}{\text{Total}}$.

Probability Mass Function A way of defining the relative frequency with a function.

1.2 Properties of Probability

A **set** is a group or collection of well-defined objects. The set of all possible outcomes is called the **sample space** and is denoted by Ω or S.

Operations of Sets

Union $A \cup B = \{x \mid x \in A \;\&\; x \in B\}$

Intersection $A \cap B = \{x \mid x \in A \text{ or } x \in B\}$

Complement $\bar{A}$ or $A^C = \{x \mid x \in A \;\&\; x \not\in A\}$

Special Sets

Universal Set Set of all events, denoted by U.

Empty Set Nothing in this set, denoted by $\emptyset$.

Singleton A set with one element.

Mutually Exclusive Sets $A \cap B = \emptyset$

Exhaustive Sets Mutually exclusive but the union covers U.

Equally Likely Events Events that have the same probability of happening.

1.3 Methods of Enumeration

Multiplication Principle

If one experiment has n_1 outcomes, and another experiment has n_2 outcomes, the composite experiment will have $n_1 n_2$ outcomes.

Permutations

Now suppose that that n positions are to be filled with n objects. There are $(n)(n-1)\ldots(2)(1)$ or $n!$ possible arrangements. Each of these arrangements is called a **permutation**. A permutation of n objects with r taken at a time is denoted as follows.

$${}_n\mathrm{P}_r = \frac{n!}{(n-r)!}$$

Combinations

The number of ways in which r objects can be selected without replacement from n objects is called a **combination**. A combination of n objects with r taken at a time is denoted as follows.

$$\binom{n}{r} = {}_n\mathrm{C}_r = \frac{n!}{r!\,(n-r)!}$$

1.4 Conditional Probability

The **conditional probability** of an event, or the probability of A occurring given that B will occur is denoted as follows.

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

By rearranging the above equation, the probability of both A and B occuring can be denoted as follows.

$$P(A \cap B) = P(A \mid B) * P(B)$$

$$P(A \cap B) = P(B \mid A) * P(A)$$

1.5 Independent Events

If the occurrence of one event will not change the probability of the occurrence of another event, these events are **independent**. Properties of independent events are as follows.

$$P(B \mid A) = P(B) \qquad P(A \mid B) = P(A)$$

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

If an event does not satisfy these properties, the events are called **dependent**. Below are other sets that are independent if A and B are independent.

$$A \text{ and } B' \qquad A' \text{ and } B \qquad A' \text{ and } B'$$

1.6 Bayes's Theorem

The probability of an event A can be calculated as follows.

$$P(A) = \sum_{i=1}^{m} P(B_i \cap A) = \sum_{i=1}^{m} P(B_i)\,P(A \mid B_i)$$

To determine the probability of a specific outcome B_k, the equation for conditional probability can be used to obtain **Bayes's Theorem**, which is as follows.

$$P(B_k \mid A) = \frac{P(B_k)P(A \mid B_k)}{\sum_{i=1}^{m} P(B_i)\,P(A \mid B_i)} \qquad \text{for } k = 1, 2, 3 \ldots m$$

2 Discrete Distributions

2.1 Random Variables of the Discrete Type

Given a random experiment with outcome space S, a function X that assigns one real number to each element int S is called a **random variable**. If S contains a finite number of points, or the points of S can be put into a one-to-one correspondence with positive integers, X becomes a random variable of the **discrete type**. A probability mass function of a discrete random variable satisfies the following properties.

$$f(x) > 0 \qquad \sum_{x \in S} f(x) = 1 \qquad P(X \in A) = \sum_{x \in A} f(x)$$

There are two ways of representing a probability mass function. One way is with a **bar graph**, in which $f(x)$ is represented with a vertical line segment. Another way of representing a probability mass function is with a **probability histogram**, where each probability is depicted by a rectangle with base length 1.

2.2 Mathematical Expectation

There are many important characteristics of a discrete distribution. One is the **mathematical expectation**. The mathematical expectation of a function of X is as follows.

$$E[X] = \sum_{x \in S} x\, f(x)$$

Below are some properties of the expected value of a function, E.

$$E(c) = c \qquad E[cX] = cE[X]$$

$$E[c_1 X + c_2 X] = c_1 E[X] + c_2 E[X]$$

2.3 The Mean, Variance, and Standard Deviation

The **mean** of a random variable X is as follows.

$$\mu = \sum_{x \in S} x\, f(x) = E[X]$$

The **variance** of the random variable X, or of the distribution of X is as follows.

$$\sigma^2 = \sum_{x \in S} (x-\mu)^2\, f(x) = E([(X-\mu)]^2 = E[X^2] - \mu^2$$

The **standard deviation** is the square root of the variance.

$$\sigma = \sqrt{\sigma^2}$$

The r^{th} **moment** of the distribution about b is

$$E[(X-b)^r] = \sum_{x \in S} (x-b)^r f(x)$$

The r^{th} **factorial moment** is

$$E[(X)_r] = E[X(X-1)(X-2)\ldots(X-r+1)]$$

2.4 Bernoulli Trials and the Binomial Distribution

A **Bernoulli experiment** is a random experiment, the outcome of which can be classified as a success or a failure. A sequence of **Bernoulli trials** occurs when a Bernoulli experiment is performed several independent times so that the probability of success reminds the same from trial to trial. Let the probability of success be p and the probability of failure be $q = 1-p$.

If there are only outcomes of success and failure, we can say that X has a **Bernoulli distribution**. The expected value of X can be written as follows.

$$\mu = E[X] = \sum_{x=0}^{1} x p^x (1-p)^{1-x} = p$$

The variance of X can be written as follows.

$$\sigma^2 = \sum_{x=0}^{1} (x-p)^2 p^x (1-p)^{1-x} = \sqrt{pq}$$

If the trials are independent, and the probabilities of success and failure on each trial are p and q respectively, X has a **binomial distribution**, and the probability mass function of X is as follows.

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

2.5 The Moment-Generating Function

Let X be a discrete random variable with probability mass function $f(x)$ and space S. If there is a positive number h such that $-h \leq t \leq h$, the **Moment-Generating function** can be defined as follows.

$$E[e^{tX}] = \sum_{x \in S} e^{tx} f(x)$$

The Moment-Generating function can be related to the mean and variance of a distribution.

$$\mu = M'(0) = E[X] \qquad \sigma^2 = M''(0) - [M'(0)]^2 = E[X^2] - E[X]^2$$

A **negative binomial distribution** has a probability mass function defined as follows.

$$f(x) = \binom{x-1}{r-1} p^r \, q^{x-r}$$

If a negative binomial distribution has an r value of 1, the distribution is said to be a **geometric distribution** which has a probability mass function as follows.

$$f(x) = p(1-p)^{x-1}$$

2.6 The Poisson Distribution

Let the number of changes that occur over a continuous interval be counted. We can then define an **approximate Poisson process** with a parameter $\lambda > 0$ if the following conditions are satified.

- The number of changes occuring on nonoverlapping intervals are independent.
- The probability of exactly one change occuring in a sufficiently short interval of length h is about λh.
- The probability of two or more changes occuring in a sufficiently short interval is essentially zero.

We say that the random variable X has a **Poisson distribution** if its probability mass function is of the form

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

The mean (μ) and variance (σ^2) of a Poisson distribution are both equal to λ.

3 Continuous Distributions

3.1 Continuous-Type Data

Many experiments do not have integers as outcomes. If the measurements could come from an interval of possible outcomes, this data would be of the **continuous type**. For this type of data, group the data into classes and construct a histogram from these classes. The interval with the largest class height is called the **modal class** and the respective class mark is called the **mode**. There's more here but it wasn't really covered that much in class.

3.2 Exploratory Data Analysis

This really wasn't covered much either, but some notable things are the idea of **percentiles**. The 50th percentile of a distribution is known as the **median**, the 25th percentile of a distribution is known as the **first quartile**, and the 75th percentile of a distribution is known as the **third quartile**. The distance between the first and third quartile is called the **interquartile range (IQR)**. You can make something called a **box plot** summarizing the minimum, first quartile, median, third quartile, and maximum. This is called a **five-number summary**.

3.3 Random Variables of the Continuous Type

The relative frequency histogram $h(x)$ associated with n observations of a random variable of the continuous type is a nonnegative function defined so that the total area between its graph and the x-axis equals 1. The probability over an interval given as $P(a < X < b)$ is given as

$$\int_a^b f(x)\,dx$$

This means that the probability is the area bounded by the graph of $f(x)$, the x-axis, and the lines $x = a$ and $x = b$. The **probability density function** of a random variable X of the **continuous type** with space S that is an interval or a union of integrals, is an integrable function of $f(x)$ satisfying the following conditions.

- $f(x) > 0 \qquad x \in S$
- $\int_S f(x)\,dx = 1$
- $P(a < X < b) = \int_a^b f(x)\,dx$

The **cumulative distribution function** of a random variable X of continuous type is as follows.

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

The mean, variance, and the moment-generating function of a continuous variable with a probability density function $f(x)$.

$$\mu = E[X] = \int_{-\infty}^{\infty} x\,f(x)\,dx$$

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2\,f(x)\,dx$$

$$M(t) = \int_{-\infty}^{\infty} e^{tx}\,f(x)\,dx$$

3.4 The Uniform and Exponential Distributions

The random variable X has a **uniform distribution** if its probability density function is equal to a constant on its support. If the support is on the interval $[a, b]$, then

$$f(x) = \frac{1}{b-a} \qquad a \leq x \leq b$$

The mean, variance and moment-generating function of X are as follows.

$$\mu = \frac{a+b}{2} \qquad \sigma^2 = \frac{(b-a)^2}{12} \qquad M(t) = \begin{cases} \frac{e^{tb}-e^{ta}}{t(b-a)} & t \neq 0 \\ 1 & t = 0 \end{cases}$$

Sometimes in the Poisson process, we let $\lambda = \frac{1}{\theta}$ and we say that random variable X has an **exponential distribution** if its probability density function is defined by

$$f(x) = \frac{1}{\theta} e^{\frac{-x}{\theta}}$$

The moment-generating function, mean, and variance of X are as follows.

$$M(t) = \int_0^\infty e^{tx} \left(\frac{1}{\theta}\right) e^{-\frac{x}{\theta}} \, dx$$

$$\mu = M'(0) = \theta \qquad \sigma^2 = M''(0) - [M'(0)]^2 = \theta^2$$

3.5 The Gamma and Chi-Squared Distributions

Chi-Squared distributions were not covered in class, so I'm only doing gamma. The **gamma function** is defined by

$$\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} \, dy$$

The random variable X has a **gamma distribution** if its probability density function is defined by

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}}$$

The moment-generating function, mean, and variance are

$$M(t) = \frac{1}{(1-\theta t)^\alpha} \qquad \mu = \alpha\theta \qquad \sigma^2 = \alpha\theta^2 = \frac{1}{1-\theta t}$$

3.6 The Normal Distribution

The random variable X has a **normal distribution** if it's probability density function is

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

The moment-generating function, mean, and variance are as follows.

$$M(t) = \exp\left[\mu t + \frac{\sigma^2 t^2}{2}\right]$$

$$\mu = E[X] = \mu$$

$$\sigma^2 = M''(0) - [M'(0)]^2 = \sigma^2$$