

STAT 230
Midterm II Notes

Sami Yousef

Contents

Chapter 1	Discrete and random variables	Page 2
1.1	Random Variables and Probability Functions	2
1.2	Discrete Uniform Distribution	2
1.3	Hypergeometric Distribution	3
1.4	Binomial Distribution	4
1.5	Negative Binomial Distribution	5
1.6	Geometric Distribution	6
1.7	Poisson Distribution (from Binomial)	6
1.8	Poisson Distribution (from Poisson Process)	7
Chapter 2	Expected Value and Variance	Page 8
2.1	Summarizing Data on Random Variables	8
2.2	Expectation of a Random Variable	8
2.3	Applications of Expectation	9
2.4	Means of Distributions	9
2.5	Variances of Distributions	10
Chapter 3	Continuous Random Variables	Page 12
3.1	General Terminology and Notation	12
3.2	Continuous Uniform Distribution	13

Chapter 1

Discrete and random variables

1.1 Random Variables and Probability Functions

Definition 1.1.1: Random Variable (aka "RV" or "rv")

A function that maps a sample space to the real numbers.

Definition 1.1.2: Range of a random variable

The set of all possible values that the random variable can take. We can denote the range of X by $X(S)$, where S is the sample space.

Definition 1.1.3: Discrete random variable

A random variable that takes on a finite or countably infinite number of values.

Definition 1.1.4: Continuous random variable

A random variable that takes on an uncountably infinite number of values.

Note:-

A random variable can be both discrete and infinite. For example, consider X = the number of goals scored in a season. The range of X is $\{0, 1, 2, \dots\}$, which is both infinite and discrete!

Example 1.1.1 (Range of a random variable)

If X is the number of calls to a help line in a day, then the range of X is $\{0, 1, 2, \dots\}$.

If X is the number of heads in 10 coin flips, then the range of X is $\{0, 1, 2, \dots, 10\}$.

If X is the duration spent studying for STAT 230, then the range of X is $(0, \infty)$.

I'm going to skip the rest of the section because it's easy and boring. Also, I just found out it's not on the MT lol.

1.2 Discrete Uniform Distribution

Definition 1.2.1: Discrete Uniform Distribution

If a rv X has a range $\{a, a + 1, \dots, b\}$ with each value being equally likely, then X has a discrete uniform distribution on a to b . Denoted by $X \sim U(a, b)$.

Example 1.2.1 (Discrete Uniform Distribution)

If X is the number of heads in 10 coin flips, then $X \sim U(0, 10)$.

Question 1

Suppose $X \sim (a, b)$. Find $f(x)$ and $F(x)$.

Solution: We know that $\sum f(x) = 1$. Since each value is equally likely, then $|A|f(x) = 1$, where $|A|$ is the size of the range of X . Therefore, $f(x) = \frac{1}{b-a+1}$, for $x = a, a+1, \dots, b$.

For $F(x)$, we know that $F(x) = P(X \leq x)$.

For $x \leq a$, $F(x) = 0$. For $x \geq b$, $F(x) = 1$.

For $a < x < b$, $F(x) = \sum_{i=a}^{\lfloor x \rfloor} f(i) = (\lfloor x \rfloor - a + 1)f(x) = \frac{\lfloor x \rfloor - a + 1}{b - a + 1}$.

Question 2

Suppose three 6 sided dice are rolled. Let X denote the largest number that is rolled. Compute $f(x)$.

Solution: Let Y_i denote the number rolled on the i th die. Then $Y_i \sim U(1, 6)$, for $i = 1, 2, 3$.

Now consider $F(x)$. Since X is the maximum of Y_i , we know $X \leq x \iff Y_i \leq x$ for all $i \in 1, 2, 3$.

So, $F(x) = P(X \leq x) = P(Y_1 \leq x \wedge Y_2 \leq x \wedge Y_3 \leq x)$.

$$\begin{aligned} F(x) &= P(Y_1 \leq x \wedge Y_2 \leq x \wedge Y_3 \leq x) \\ &= \frac{\lfloor x \rfloor}{6} \cdot \frac{\lfloor x \rfloor}{6} \cdot \frac{\lfloor x \rfloor}{6} \\ &= \frac{\lfloor x \rfloor^3}{216} \end{aligned}$$

Now we can compute $f(x)$ since we know $f(x) = F(x) - F(x-1)$.

$$\begin{aligned} f(x) &= F(x) - F(x-1) \\ &= \frac{\lfloor x \rfloor^3}{216} - \frac{\lfloor x-1 \rfloor^3}{216} \\ &= \frac{3\lfloor x \rfloor^2 - 3\lfloor x \rfloor + 1}{216} \end{aligned}$$

1.3 Hypergeometric Distribution

Definition 1.3.1: Hypergeometric Distribution

Consider a population that consists of N objects, of which r are considered “successes” and the remaining $N - r$ are considered “failures”.

Suppose that n objects are drawn at random from the population without replacement. Let X be the number of successes in the n objects drawn. Then X has a hypergeometric distribution with parameters N , r , and n . Denoted by $X \sim \text{hyp}(N, r, n)$.

$$f(x)_X = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}, \text{ for } \max(0, n+r-N) \leq x \leq \min(r, n).$$

Example 1.3.1 (Hypergeometric Distribution)

Suppose X is the number of aces in a hand of 5 cards drawn from a standard deck of 52 cards. Then $X \sim \text{hyp}(52, 4, 5)$.

Question 3

Consider drawing a 5 card hand at random from a standard 52 card deck of playing cards. What is the probability that the hand contains at least 3 K's

Solution: Let X be the number of K's in the hand.
Then $X \sim \text{hyp}(52, 4, 5)$. So we have

$$\begin{aligned} P(X \geq 3) &= P(X = 3) + P(X = 4) \\ &= \frac{\binom{4}{3} \binom{48}{2}}{\binom{52}{5}} + \frac{\binom{4}{4} \binom{48}{1}}{\binom{52}{5}} \\ &= 0.00175 \end{aligned}$$

Example 1.3.2 (Hypergeometric Distribution)

Suppose a bag contains 20 balls of which 4 are blue, 6 are red and 10 are green. We draw 5 balls at random without replacement. Let X denote the number balls drawn that are not blue. Then $X \sim \text{hyp}(20, 16, 5)$.

1.4 Binomial Distribution

Definition 1.4.1: Bernoulli Trial

A trial that has only two possible outcomes, success and failure.

Definition 1.4.2: Binomial Distribution

Let X be the number of successes in n independent Bernoulli trials, each with probability p of success. Then X has a binomial distribution with parameters n and p . Denoted by $X \sim \text{Binomial}(n, p)$.

$$f(x)_X = \binom{n}{x} p^x (1-p)^{n-x}, \text{ for } x \in \{0, 1, \dots, n\}.$$

Example 1.4.1 (Binomial Distribution)

Suppose X is the number of heads in 10 coin flips. Then $X \sim \text{Binomial}(10, 0.5)$.

Note:-

A key assumption of the Binomial Distribution is that the trials are independent and the probability of each trial is constant. If these assumptions are not met, the Binomial Distribution would not be appropriate.

Question 4

Suppose a tack is flipped 10 times, with a probability of 0.6 of landing point up. What is the probability that the tack lands pointing up at least 3 times?

Solution: Let X be the number of times the tack lands pointing up. Then $X \sim \text{Binomial}(10, 0.6)$. We need to compute $P(X \geq 3)$.

$$\begin{aligned} P(X \geq 3) &= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &= 1 - \binom{10}{0} 0.6^0 (1-0.6)^{10-0} - \binom{10}{1} 0.6^1 (1-0.6)^{10-1} - \binom{10}{2} 0.6^2 (1-0.6)^{10-2} \\ &= 0.988 \end{aligned}$$

Theorem 1.4.1 Binomial approximation to the hypergeometric distribution

If r and N are large relative to n , and $\frac{r}{N} = p$ with $p \in [0, 1]$; then if $X \sim \text{hyp}(N, r, n)$ and $Y \sim \text{Binomial}(n, p)$, then

$$P(X = k) \approx P(Y = k).$$

Question 5: Tim Hortons

For the Roll Up the Rim promotion, Tim Hortons prints 16,225,200 cups, 2,704,200 of which offer a free coffee or donut. You buy 15 cups during the promotion. Let X = the number of times you win a coffee or donut.

1. What is the distribution of X ?
2. If we suppose $X \sim \text{Binomial}(n, p)$, what are n and p ? What assumptions are we making?
3. What is the probability you never win?
4. What is the probability you win more than once?

Solution:

1. $X \sim \text{hyp}(16,225,200, 2,704,200, 15)$
2. We will have $n = 15$ and $p = \frac{2,704,200}{16,225,200} = \frac{1}{6}$. We are assuming that the 15 cups are drawn independently and that the probability of winning is constant.
3. We need to compute $P(X = 0)$.

$$\begin{aligned} P(X = 0) &= \binom{15}{0} \left(\frac{1}{6}\right)^0 \left(1 - \frac{1}{6}\right)^{15-0} \\ &= \left(\frac{5}{6}\right)^{15} \\ &= 0.065 \end{aligned}$$

4. We need to compute $P(X \geq 2)$.

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \left(\frac{5}{6}\right)^{15} - \binom{15}{1} \left(\frac{1}{6}\right)^1 \left(1 - \frac{1}{6}\right)^{15-1} \\ &= 0.740 \end{aligned}$$

1.5 Negative Binomial Distribution

Definition 1.5.1: Negative Binomial Distribution

Let X be the number of failures of Bernoulli trials needed before observing r successes. Then X has a negative binomial distribution with parameters k and p . Denoted by $X \sim \text{NB}(k, p)$.

$$f(x)_X = \binom{x+k-1}{k-1} p^k (1-p)^x, \text{ for } x \in \{0, 1, \dots, \infty\}.$$

Note:-

There are some key differences between the negative binomial and the binomial distributions.

1. Binomial has a fixed number of trials, while negative binomial has an infinite number of successes.
2. Binomial counts the number of successes, while negative binomial counts the number of failures.

Example 1.5.1 (Negative Binomial Distribution)

Suppose a coin is flipped until 5 heads are observed, if X is the number of tails observed, then $X \sim NB(5, 0.5)$.

1.6 Geometric Distribution

Definition 1.6.1: Geometric Distribution

Let X be the number of Bernoulli trials needed to get the first success. Then X has a geometric distribution with parameter p . Denoted by $X \sim Geo(p)$. Note that the Geometric Distribution is a special case of the Negative Binomial Distribution, with $k = 1$.

$$f(x)_X = (1 - p)^x p, \text{ for } x \in \{0, 1, 2, \dots, \infty\}.$$

Note:-

It is easy to mistake the Geometric Distribution for the Binomial Distribution since they both concern Bernoulli trials. However, Binomial concerns a fixed number of trials (n) - Geometric does not!

Question 6

You start buying Roll Up the Rim cups, where you assume the probability of victory is $\frac{1}{6}$ and independent of all previous plays. You stop buying once you win for the first time. What is the probability you will have to buy at least 6 cups?

Solution: Firstly, we define X to be the number of cups you buy before winning. Then $X \sim Geo(\frac{1}{6})$. We need to compute $P(X \geq 5)$.

Notice that we want $X \geq 5$, not $X > 5$. This is because for the 6th cup to win, we need to have 5 losing cups first.

$$\begin{aligned} P(X \geq 5) &= 1 - P(X = 0) - P(X = 1) - P(X = 2) - P(X = 3) - P(X = 4) \\ &= 1 - \frac{1}{6} \left(\frac{5}{6}\right)^0 - \frac{1}{6} \left(\frac{5}{6}\right)^1 - \frac{1}{6} \left(\frac{5}{6}\right)^2 - \frac{1}{6} \left(\frac{5}{6}\right)^3 - \frac{1}{6} \left(\frac{5}{6}\right)^4 \\ &= 0.402 \end{aligned}$$

1.7 Poisson Distribution (from Binomial)

Historically, the Binomial Distribution was difficult to compute. The Poisson Distribution was developed as a way to approximate the Binomial Distribution when n is large and p is small.

Definition 1.7.1: Approximating the Binomial Distribution

As $n \rightarrow \infty$ and $p \rightarrow 0$, the p.f. for the Binomial Distribution approaches

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

where $\lambda = np$.

Question 7

Define X to be the number of people with a certain disease. If we test 1000 people, and the probability of testing positive is 0.01, what is the probability that 5 people test positive?

Solution: We will use the Poisson Distribution to approximate the Binomial Distribution. We know $X \sim \text{Binomial}(1000, 0.01)$ or $X \sim \text{Poi}(10)$. We need to compute $P(X = 5)$.

$$\begin{aligned} P(X = 5) &= \frac{e^{-10} 10^5}{5!} \\ &= 0.0378 \end{aligned}$$

It is also possible to use the Binomial Distribution directly:

$$\begin{aligned} P(X = 5) &= \binom{1000}{5} (0.01)^5 (0.99)^{1000-5} \\ &= 0.0378 \end{aligned}$$

As you can see, the two methods give a very similar answer because n is large and p is small!

1.8 Poisson Distribution (from Poisson Process)

Definition 1.8.1: Poisson Process

A process that counts the number of events that occur in a fixed time interval and satisfies the following conditions:

1. **Independence:** The number of occurrences in non-overlapping intervals are independent
2. **Indivisuality or Singularity:** The probability of more than one occurrence in a small interval is zero
3. **Homogeneity or Uniformity:** The rate of occurrence is constant over time

We denote the number of occurrences in a time interval of length t by $X_t \sim \text{Poi}(\lambda t)$ and $f_t(x) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}$.

Example 1.8.1 (Poisson Process)

Suppose a website has 100 visitors per minute (following a Poisson Process). What is the probability that 5 visitors arrive in a 10 second interval?

$$X_{10} \sim \text{Poi}\left(\frac{100}{60} \cdot 10\right).$$

$$P(X_{10} = 5) = \frac{e^{-\frac{100}{6}} \left(\frac{100}{6}\right)^5}{5!} = 0.00061.$$

Chapter 2

Expected Value and Variance

2.1 Summarizing Data on Random Variables

Definition 2.1.1: Mean

The mean of a set of data is defined as the sum of the data points divided by the number of data points.

Definition 2.1.2: Median

The median of a set of data is the middle value of the data points.

Definition 2.1.3: Mode

The mode of a set of data is the most common value of the data points.

Example 2.1.1

Do you really need an example?

2.2 Expectation of a Random Variable

Definition 2.2.1: Expected Value

Suppose X is a random variable with p.f. $f(x)$ and range A . Then the expected value of X is defined as:

$$E(X) = \sum_{x \in A} x f(x)$$

The expected value of X is sometimes called the mean of X or the first moment of X , denoted by μ_X .

Example 2.2.1

Suppose X is the outcome of a fair die roll. What is $E(X)$?
Since X is a fair die roll, $f(x) = \frac{1}{6}$ for all $x \in \{1, 2, 3, 4, 5, 6\}$.

$$\begin{aligned} E(X) &= \sum_{x \in \{1, 2, 3, 4, 5, 6\}} x \frac{1}{6} \\ &= \frac{1}{6} \cdot (1 + 2 + 3 + 4 + 5 + 6) \\ &= \frac{21}{6} \\ &= 3.5 \end{aligned}$$

Question 8: Real World Application

Suppose it costs \$100 to enter a contest. The contest entails that you receive a random amount of money between \$1 and \$400. What is the **net** expected value of the contest?

Solution:

Let X be the amount of money you receive. Then $X \sim U(1, 400)$.
We need to compute $E(X) - 100$.

$$\begin{aligned} E(X) - 100 &= \sum_{x \in \{1, 2, \dots, 400\}} x \frac{1}{400} - 100 \\ &= \frac{1}{400} \cdot (1 + 2 + \dots + 400) - 100 \\ &= \frac{1}{400} \cdot \frac{400 \cdot 401}{2} - 100 \\ &= \frac{401}{2} - 100 \\ &= 100.5 \end{aligned}$$

The expected value is positive, so you should enter the contest!

Theorem 2.2.1 Law of The Unconscious Statistician

Let X be a r.v. with p.f. $f(x)$ and range A . The expected value of some function of x , $g(x)$ is given by:

$$E(g(X)) = \sum_{x \in A} g(x)f(x)$$

Note:-

$E(\cdot)$ is a linear operator. That is, $E(aX + b) = aE(X) + b$.

2.3 Applications of Expectation

2.4 Means of Distributions

In this section, we define the mean of distributions we looked at earlier.

Definition 2.4.1: Mean of Binomial Distribution

Let $X \sim \text{Binomial}(n, p)$. Then $E(X) = np$.

Definition 2.4.2: Mean of Poisson Distribution

Let $X \sim \text{Poi}(\lambda)$. Then $E(X) = \lambda$.

Definition 2.4.3: Mean of Hypergeometric Distribution

Let $X \sim \text{Hyper}(n, N, r)$. Then $E(X) = \frac{nr}{N}$.

Definition 2.4.4: Mean of Geometric Distribution

Let $X \sim \text{Geo}(p)$. Then $E(X) = \frac{1-p}{p}$.

Definition 2.4.5: Mean of Negative Binomial Distribution

Let $X \sim \text{NB}(k, p)$. Then $E(X) = \frac{k(1-p)}{p}$.

2.5 Variances of Distributions

Definition 2.5.1: Variance

The variance of a random variable X is defined as:

$$\text{Var}(X) = E((X - E(X))^2)$$

Another way to define variance is **the inside-outside method**:

$$\text{Var}(X) = E(X^2) - E(X)^2$$

Corollary 2.5.1 Some Properties of Variance

1. $\forall \text{ r.v. } X, \text{Var}(X) \geq 0$
2. $\forall \text{ r.v. } X, \text{Var}(X) = 0 \iff X \text{ is a constant}$
3. $E(X^2) \geq E(X)^2$
4. $\text{Var}(aX + b) = a^2 \text{Var}(X)$

Definition 2.5.2: Standard Deviation

The standard deviation of a random variable X is defined as:

$$\sigma(X) = \sqrt{\text{Var}(X)}$$

Definition 2.5.3: Variance of Binomial Distribution

Let $X \sim \text{Binomial}(n, p)$. Then $\text{Var}(X) = np(1 - p)$.

Definition 2.5.4: Variance of Poisson Distribution

Let $X \sim Poi(\lambda)$. Then $\text{Var}(X) = \lambda$.

Definition 2.5.5: Variance of Hypergeometric Distribution

Let $X \sim hyp(n, N, r)$. Then $\text{Var}(X) = \frac{nr(1-\frac{r}{N})}{N(N-1)}$.

Definition 2.5.6: Variance of Geometric Distribution

Let $X \sim Geo(p)$. Then $\text{Var}(X) = \frac{1-p}{p^2}$.

Definition 2.5.7: Variance of Negative Binomial Distribution

Let $X \sim NB(k, p)$. Then $\text{Var}(X) = \frac{k(1-p)}{p^2}$.

Chapter 3

Continuous Random Variables

3.1 General Terminology and Notation

Definition 3.1.1: Continuous Random Variable

A continuous random variable is a random variable whose range is a subset of the real numbers.

Definition 3.1.2: Probability Density Function

A continuous r.v. has a probability density function $f(x)$ that satisfies the following conditions:

1. $f(x) \geq 0$ for all $x \in \mathbb{R}$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$
3. $P(a \leq X \leq b) = \int_a^b f(x)dx$

Definition 3.1.3: The support of a c.r.v.

The support of a c.r.v., $\text{supp}(x)$, is defined as the set of all x such that $f(x) \neq 0$.

Definition 3.1.4: Cumulative Distribution Function

The cumulative distribution function of a c.r.v. X is defined as:

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

Corollary 3.1.1

When $f(x)$ is continuous, we have $f(x) = \frac{d}{dx}F(x)$.

This is useful because $F(x)$ is easier to work with than $f(x)$.

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(x)dx.$$

Definition 3.1.5: Quantile

Suppose X is a c.r.v. with CDF $F(x)$. Then the p -th quantile of X is the value of $q(p)$ such that $P(X \leq q(p)) = p$.

Theorem 3.1.1 Change of Variables

Let X be a c.r.v. with CDF $F(x)$ and $Y = g(X)$ be a c.r.v. with CDF $G(y)$. Then $G(y) = F(g^{-1}(y))$.

Recepie for Change of Variables:

1. Find the range of Y .
2. Write the CDF of Y as a function of X
3. Use $F(x)$ to find $G(y)$
4. (Optional) Find the PDF of Y by taking the derivative of $G(y)$

Example 3.1.1 (Change of Variables)

Let X be a c.r.v. with CDF $F(x) = 0.25x$ and $Y = \sqrt{X}$. Find the CDF of Y .

Solution:

1. The range of Y is $[0, \infty)$.
2. $G(y) = P(Y \leq y) = P(\sqrt{X} \leq y) = P(X \leq y^2) = F(y^2)$
3. $G(y) = F(y^2) = 0.25y^2$
4. $f(y) = \frac{d}{dy} G(y) = \frac{d}{dy} 0.25y^2 = 0.5y$ (Optional)

Definition 3.1.6: Expected Value of a c.r.v.

Let X be a c.r.v. with CDF $F(x)$. Then $E(X) = \int_{-\infty}^{\infty} x f(x) dx$.

The LOTUS theorem states that $E(G(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx$.

Other properties defined for discrete r.v. also hold for c.r.v. (e.g. linearity of $E(\cdot)$, $\text{Var}(X) = E(X^2) - E(X)^2$, etc.)

3.2 Continuous Uniform Distribution

Definition 3.2.1: Continuous Uniform Distribution

X is said to have a continuous uniform distribution on (a, b) if X takes values in (a, b) or $[a, b]$ where all subintervals of fixed length have the same probability.

This is abbreviated as $X \sim U(a, b)$.

The PDF of X is $f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$.

The CDF of X is $F(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x > b \end{cases}$.