Lecture Notes

Week 1: Revision of Probability Models

Introduction to Stochastic Methods

Randomness

Definition: Randomness
Randomness refers to the lack of pattern or predictability in events. Key properties include unpredictability, patternlessness, statistical regularity, independence, and reproducibility in aggregates.

Random Variable

Definition: Random Variable
Consider an experiment with a sample space S on which probabilities are defined. A random variable X is a function that assigns a real value to each outcome of the experiment. For any set of real numbers C, the probability that X will have a value that is contained in the set C is equal to the probability that the outcome of the experiment is contained in $X^{-1}(C)$. That is,

$$P\{X \in C\} = P\{X^{-1}(C)\},$$

where $X^{-1}(C)$ is the event consisting of all outcomes $s \in S$ such that $X(s) \in C$.

The distribution function F of the random variable X is defined for all real numbers by

$$F(x) = P\{X \leq x\} = P\{X \in (-\infty, x]\}.$$

We will use the notation $\bar{F}$ to represent $1 - F$; that is,

$$\bar{F}(x) = 1 - F(x) = P\{X > x\}.$$

Example:
Consider a dice roll. Let X be the outcome of rolling a fair six-sided dice, with $X \in \{1, 2, 3, 4, 5, 6\}$.

Probability Distribution

Definition: Probability Distribution
A probability distribution describes how probabilities are distributed over the values of a random variable, providing the probabilities of occurrence of different possible outcomes.

Examples of Probability Distributions

Example: Geometric Distribution
Describes the number of trials needed to get the first success in a sequence of independent Bernoulli trials. The PMF is given by:

$$P(X = k) = (1 - p)^{k-1} p,$$

where k is the number of trials until the first success, and p is the probability of success on each trial.
Example: The probability of getting the first head in a series of coin flips.
Example: Binomial Distribution
Gives the probability of observing a specific number of successes in a fixed number of independent Bernoulli trials. The PMF is:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k},$$

where n is the number of trials, k is the number of successes, and p is the probability of success on each trial.
Example: The probability of getting exactly three heads in five tosses of a fair coin.
Example: Poisson Distribution
Describes the probability of a given number of events happening in a fixed interval of time or space. The PMF is:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!},$$

where k is the number of events, λ is the average number of events per interval.
Example: The number of cars passing through a checkpoint in an hour.
Example: Exponential Distribution
Describes the time between events in a Poisson point process. The PDF is:

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

for $x \geq 0$, where λ is the rate parameter.
Example: The amount of time until the next earthquake occurs in a given region.
Example: Normal Distribution
Describes how the values of a variable are distributed. The PDF is:

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

where μ is the mean and σ is the standard deviation.
Example: The distribution of heights of adult men in a specific population.

Expectation

Definition: Expectation
The expectation (or expected value) of a random variable is a measure of the central tendency of its probability distribution. It is denoted by $E[X]$ for a random variable X and is calculated as:

$$E[X] = \sum_x x P(X = x)$$

for discrete random variables, and

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

for continuous random variables, where $P(X = x)$ is the probability mass function for discrete variables and $f(x)$ is the probability density function for continuous variables.

Variance

Definition: Variance
Variance measures the dispersion of a random variable's values around its mean. It is denoted by $\mathrm{Var}(X)$ or σ_X^2 and is calculated as:

$$\mathrm{Var}(X) = E[(X-\mu)^2] = E[X^2] - (E[X])^2,$$

where $\mu = E[X]$ is the mean (or expectation) of X.

Detailed Calculations for the Binomial Distribution

Let's calculate the expectation and variance of a binomial distribution with parameters n and p, where n is the number of trials and p is the probability of success.
Expectation:

$$E[X] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}.$$

We use the binomial theorem and properties of binomial coefficients to simplify this expression, recognizing that this is equivalent to np by considering the derivative of the binomial expansion $(p + (1-p))^n$.
Variance:

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2.$$

First, calculate $E[X^2]$:

$$E[X^2] = \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k}.$$

This requires using the binomial theorem and properties of binomial coefficients, similar to the expectation but involves a more complex manipulation. Eventually, we find that $E[X^2] = np(1-p) + n^2p^2$. Substituting $E[X] = np$ into the variance formula gives:

$$\mathrm{Var}(X) = np(1-p) + n^2p^2 - (np)^2 = np(1-p).$$

This section not only adds the moment-generating function and its application but also provides a detailed walkthrough of calculating expectation and variance for the binomial distribution, showcasing the mathematical reasoning behind these fundamental concepts.

Moments

Definition: Moments
Moments are quantitative measures used to describe the shape of a probability distribution. The n-th moment of a random variable X about the mean is defined as:

$$\mu_n = E[(X-\mu)^n],$$

where $\mu = E[X]$ is the mean of X. The first moment about the mean is the mean itself, the second moment about the mean is the variance, and higher moments describe skewness, kurtosis, and other aspects of the distribution's shape.

Moment-Generating Function

Definition: Moment-Generating Function
The moment-generating function (MGF) of a random variable X is defined as:

$$M_X(t) = E[e^{tX}],$$

where t is a real number, and the expectation is taken over the probability distribution of X. The MGF, if it exists, uniquely determines the probability distribution of X and can be used to find all the moments of the distribution since the n-th moment of X is given by the n-th derivative of $M_X(t)$ evaluated at $t = 0$:

$$\mu_n = \left. \frac{d^n M_X(t)}{dt^n} \right|_{t=0}.$$

Example:
Consider a random variable X that follows an exponential distribution with rate parameter λ. The MGF of X is:

$$M_X(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t},$$

for $t < \lambda$. The first derivative of $M_X(t)$ with respect to t, evaluated at $t = 0$, gives the mean (the first moment) of X:

$$E[X] = \left. \frac{dM_X(t)}{dt} \right|_{t=0} = \frac{1}{\lambda}.$$

Exercise: Moment-Generating Function of a Discrete Uniform Distribution

Consider a discrete uniform random variable X that can take on the values $1, 2, 3, \ldots, n$ with equal probability.

Task: Find the moment-generating function (MGF) of X, $M_X(t)$, and use it to compute the first and second moments ($E[X]$ and $E[X^2]$).

Solution

The MGF of a random variable X is defined as $M_X(t) = E[e^{tX}]$. For a discrete uniform distribution over $1, 2, \ldots, n$, this becomes:

$$M_X(t) = \frac{1}{n} \sum_{k=1}^{n} e^{tk} = \frac{1}{n} \frac{e^t(1 - e^{tn})}{1 - e^t}, \quad \text{for } t \neq 0.$$

To find the first and second moments, we differentiate $M_X(t)$ with respect to t and evaluate at $t = 0$:

- First moment ($E[X]$): $M'_X(0) = \frac{1}{2}(n + 1)$.
- Second moment ($E[X^2]$): $M''_X(0) = \frac{1}{6} n(n + 1)(2n + 1)$.

Exercise: Moment-Generating Function of a Bernoulli Distribution

Consider a Bernoulli random variable Y with parameter p, which represents the probability of success ($Y = 1$).

Task: Find the moment-generating function (MGF) of Y, $M_Y(t)$, and use it to compute the first moment ($E[Y]$).

Solution

The MGF of Y is given by:

$$M_Y(t) = E[e^{tY}] = pe^t + (1-p).$$

To find the first moment ($E[Y]$), we differentiate $M_Y(t)$ with respect to t and evaluate at $t = 0$:

- First moment ($E[Y]$): $M'_Y(0) = p$.

Conditional Probability

Recall that for any two events E and F, the conditional probability of E given F is defined, as long as $P(F) > 0$, by

$$P(E|F) = \frac{P(EF)}{P(F)}$$

Hence, if X and Y are discrete random variables, then it is natural to define the conditional probability mass function of X given that $Y = y$, by

$$p_{X|Y}(x|y) = P\{X = x|Y = y\} = \frac{P\{X = x, Y = y\}}{P\{Y = y\}} = \frac{p(x,y)}{p_Y(y)}$$

where $p(x, y)$ is the joint probability mass function of X and Y, and $p_Y(y)$ is the marginal probability mass function of Y.

Let E denote an arbitrary event and define the indicator random variable X by

$$X = \begin{cases} 1, & \text{if } E \text{ occurs,} \\ 0, & \text{if } E \text{ does not occur.} \end{cases}$$

It follows from the definition of X that

$$E[X] = P(E),$$

$$E[X|Y = y] = P(E|Y = y),$$

for any random variable Y.

Therefore, from Equations (3.3a) and (3.3b) we obtain

$$P(E) = \sum_y P(E|Y = y)P(Y = y), \text{ if } Y \text{ is discrete}$$

$$P(E) = \int_{-\infty}^{\infty} P(E|Y = y) f_Y(y)\, dy, \text{ if } Y \text{ is continuous}$$

where $f_Y(y)$ is the probability density function of Y.

Example

If X_1 and X_2 are independent binomial random variables with respective parameters (n_1, p) and (n_2, p), calculate the conditional probability mass function of X_1 given that $X_1 + X_2 = m$.

Solution: With $q = 1 - p$,

$$
\begin{aligned}
P\{X_1 = k | X_1 + X_2 = m\} &= \frac{P\{X_1 = k, X_1 + X_2 = m\}}{P\{X_1 + X_2 = m\}} \\
&= \frac{P\{X_1 = k, X_2 = m - k\}}{P\{X_1 + X_2 = m\}} \\
&= \frac{P\{X_1 = k\} P\{X_2 = m - k\}}{P\{X_1 + X_2 = m\}} \\
&= \frac{\binom{n_1}{k} p^k q^{n_1 - k} \binom{n_2}{m-k} p^{m-k} q^{n_2 - (m-k)}}{\binom{n_1+n_2}{m} p^m q^{n_1+n_2-m}}
\end{aligned}
$$

where we have used that $X_1 + X_2$ is a binomial random variable with parameters $(n_1 + n_2, p)$ (see Example 2.44). Thus, the conditional probability mass function of X_1, given that $X_1 + X_2 = m$, is

$$
P\{X_1 = k | X_1 + X_2 = m\} = \frac{\binom{n_1}{k}\binom{n_2}{m-k}}{\binom{n_1+n_2}{m}}
$$

The distribution given by Equation (3.1), first seen in Example 2.34, is known as the hypergeometric distribution. It is the distribution of the number of blue balls that are chosen when a sample of m balls is randomly chosen from an urn that contains n_1 blue and n_2 red balls. (To intuitively see why the balls that are chosen if X_1 and X_2 are independent is randomly chosen from the urn that contains n_1 blue and n_2 red balls.)

Example

If X and Y are independent Poisson random variables with respective means λ_1 and λ_2, calculate the conditional expected value of X given that $X + Y = n$.

Solution: Let us first calculate the conditional probability mass function of X given that $X + Y = n$. We obtain

$$
\begin{aligned}
P\{X = k | X + Y = n\} &= \frac{P\{X = k, X + Y = n\}}{P\{X + Y = n\}} \\
&= \frac{P\{X = k, Y = n - k\}}{P\{X + Y = n\}} \\
&= \frac{P\{X = k\} P\{Y = n - k\}}{P\{X + Y = n\}} \\
&= \frac{e^{-\lambda_1} \frac{\lambda_1^k}{k!} e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!}}{e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^n}{n!}} \\
&= \frac{n!}{(n-k)!k!} \frac{\lambda_1^k \lambda_2^{n-k}}{(\lambda_1 + \lambda_2)^n} \\
&= \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-k}
\end{aligned}
$$

In other words, the conditional distribution of X given that $X + Y = n$, is the binomial distribution with parameters n and $\frac{\lambda_1}{\lambda_1+\lambda_2}$. Hence,

$$
E\{X | X + Y = n\} = n \frac{\lambda_1}{\lambda_1 + \lambda_2}
$$

This result demonstrates the conditional expectation of X given the sum $X+Y$, showing that it follows a binomial distribution in the context of the given Poisson random variables.

Example

Suppose the joint density of X and Y is given by

$$f(x,y) = \begin{cases} 6xy(2-x-y), & 0<x<1, 0<y<1, \\ 0, & \text{otherwise.} \end{cases}$$

Compute the conditional expectation of X given that $Y=y$, where $0<y<1$.
Solution: We first compute the conditional density

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

$$= \frac{6xy(2-x-y)}{\int_0^1 6xy(2-x-y)\,dx}$$

$$= \frac{6xy(2-x-y)}{y(4-3y)}$$

$$= \frac{6x(2-x-y)}{4-3y}$$

Hence,

$$E[X|Y=y] = \int_0^1 x \cdot \frac{6x(2-x-y)}{4-3y}\,dx$$

$$= \frac{(2-y)^2 - \frac{6}{4}}{4-3y}$$

$$= \frac{5-4y}{8-6y}$$

This result provides the conditional expectation $E[X|Y=y]$, which is a function of y.

Exercises

Exercise 1: Understanding Randomness

1. Q1: Give two examples of random processes in real life and describe how they exhibit unpredictability and statistical regularity.

2. Q2: Discuss how randomness plays a role in simulations and modeling. Provide an example where randomness is crucial to obtaining realistic results.

Exercise 2: Random Variables

1. Q1: Define a random variable for the experiment of flipping a fair coin three times. Describe the sample space and assign a numerical value for each possible outcome.

2. Q2: For the random variable X representing the sum of numbers obtained when rolling two six-sided dice, calculate $P\{X=7\}$ and $P\{X>4\}$.

Exercise 3: Probability Distributions

1. Q1: Given a geometric distribution with $p = 0.25$, calculate the probability of obtaining the first success on the 4th trial.

2. Q2: For a binomial distribution with parameters $n = 5$ and $p = 0.5$, find the probability of obtaining exactly 3 successes.

Exercise 4: Expectation and Variance

1. Q1: Calculate the expectation and variance of a binomial distribution with $n = 10$ and $p = 0.3$.

2. Q2: A random variable X follows a Poisson distribution with $\lambda = 4$. Find the expected value and variance of X.

Exercise 5: Moment-Generating Functions

1. Q1: Derive the moment-generating function for a Bernoulli random variable with probability of success p.

2. Q2: Given the MGF of a random variable X, $M_X(t) = e^{2(e^t-1)}$, determine the probability distribution of X and calculate the first and second moments.

Exercise 6: Practical Application

1. Q1: Consider a small business that experiences a random number of customer arrivals per day, which can be modeled by a Poisson distribution with an average rate of $\lambda = 10$ customers per day. Calculate the probability of receiving more than 12 customers in a day.

2. Q2: A quality control engineer measures the diameter of a ball bearing produced by a machine. The diameters are normally distributed with a mean of 5 mm and a standard deviation of 0.1 mm. Calculate the probability that a randomly selected ball bearing has a diameter of more than 5.1 mm.

Exercise 7: Deep Dive into Distributions

1. Task: Choose one of the probability distributions discussed (e.g., geometric, binomial, Poisson, exponential, normal) and perform a detailed analysis:

   (a) Derive its mean and variance from first principles.

   (b) Sketch its probability mass function (PMF) or probability density function (PDF) for different parameter values.

   (c) Discuss its applications and limitations in real-world scenarios.

Solutions

Excercise 1

Q1: Two examples of random processes in real life are:

1. The decay of a radioactive atom. It is impossible to predict exactly when a specific atom will decay, but statistically, a large number of such atoms will decay at a rate described by a known half-life.

2. The outcome of a lottery draw. Each draw is unpredictable and independent of previous draws, yet when considering a large number of draws, patterns and probabilities become statistically discernible.

Q2: Randomness is integral to simulations and modeling as it allows for the representation of uncertain or variable factors. For instance, in climate modeling, the inherent unpredictability of weather elements, such as temperature and precipitation, is simulated using random variables to produce results that approximate the variability observed in the real climate system.

Excercise 2

Q1: A random variable for the experiment of flipping a fair coin three times can be defined as $X : \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}, \mathrm{HTT}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\} \to \{0, 1, 2, 3\}$ where X assigns a numerical value to each possible outcome representing the number of heads.

Q2: For the random variable X representing the sum of numbers obtained when rolling two six-sided dice:

- The probability of getting a sum of 7 is $P(X = 7) = \frac{1}{6}$.
- The probability of getting a sum greater than 4 is $P(X > 4) = \frac{5}{6}$.

Excercise 3

Q1: For a geometric distribution with probability $p = 0.25$, the probability of the first success on the 4th trial is given by:

$$P(X = 4) = (1 - p)^3 \times p = (0.75)^3 \times 0.25$$

Q2: For a binomial distribution with parameters $n = 5$ and $p = 0.5$, the probability of obtaining exactly 3 successes is given by the binomial probability formula:

$$P(X = 3) = \binom{n}{k} p^k (1 - p)^{n-k} = \binom{5}{3} (0.5)^3 (0.5)^2$$

Excercise 4

Q1: The expectation and variance of a binomial distribution with $n = 10$ and $p = 0.3$ are:

Expectation:

$$E[X] = n \times p = 10 \times 0.3$$

Variance:

$$\mathrm{Var}(X) = n \times p \times (1 - p) = 10 \times 0.3 \times (1 - 0.3)$$

Q2: For a Poisson distribution with $\lambda = 4$:

Expectation: $E[X] = \lambda = 4$

Variance: $\mathrm{Var}(X) = \lambda = 4$

Excercise 5

Q1: The moment-generating function (MGF) for a Bernoulli random variable with probability of success p is:

$$M_X(t) = E[e^{tX}] = pe^t + (1-p)$$

Q2: Given the MGF of a random variable X, $M_X(t) = e^{2(e^t-1)}$, the first and second moments are found by taking the first and second derivatives of $M_X(t)$ at $t = 0$.

Excercise 6

Q1: For a Poisson distribution with $\lambda = 10$, the probability of receiving more than 12 customers is:

$$P(X > 12) = 1 - \sum_{k=0}^{12} \frac{e^{-\lambda}\lambda^k}{k!}$$

Q2: For a normal distribution with mean $\mu = 5$ mm and standard deviation $\sigma = 0.1$ mm, the probability that a randomly selected ball bearing has a diameter of more than 5.1 mm is:

$$P(X > 5.1) = 1 - \Phi\left(\frac{5.1-\mu}{\sigma}\right)$$

Excercise 7

(a) For the Poisson distribution with parameter λ, the mean and variance are both equal to λ, which can be derived from the probability mass function (PMF):

$$\text{Mean} = E[X] = \lambda$$

$$\text{Variance} = \text{Var}(X) = \lambda$$

(b) The probability mass function (PMF) for a Poisson distribution is:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

The PMF can be sketched for different values of λ, which will show the distribution of probabilities for different numbers of occurrences.

(c) Applications of the Poisson distribution include modeling the number of events in a fixed interval of time or space under the assumption that these events occur with a known constant rate and independently of the time since the last event. Limitations include the assumption of independence and a constant rate, which may not hold in all real-world scenarios.