Midterm I Review

MATH 391

October 1, 2014

Logistics

- 50 minutes in class
- Please bring a calculator or smart phone with calculator app
- One 8.5×11 single-sided cheat sheet
- R will be on the exam. No need to code directly, but I present you with code and you need to interpret what probabilistic concept is being illustrated.
- Office hours on Thursday 10/2 from 2:30-4pm.

Lec 1: Experiments, Sample space, Events

- Experiments
- Sample space and its outcomes
- Events

Carefully label events of interest.

Lec 1: Kolmogorov Axioms

From *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Let S be a sample space and let E be an event in S, then

1. $0 \leq \Pr(E)$ for every $E \in S$
2. $\Pr(S) = 1$
3. Countable Additivity: Let $E_1, E_2, \ldots$ be mutually exclusive[1] events, then

$$\Pr\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \Pr\left(E_i\right)$$

[1] $E_i \cap E_j = \emptyset$ for any $i \neq j$. AKA disjoint

Lec 1: Useful Consequences of the Kolmogorov Axioms

From

1. $\Pr(E^c) = 1 - \Pr(E)$
2. $\Pr(\emptyset) = 0$
3. If $E \subset F$, then $\Pr(E) \leq \Pr(F)$
4. $\Pr(E \cap F^c) = \Pr(E) - \Pr(E \cap F)$
5. For any two events E and F (not necessarily disjoint), we have $\Pr(E \cup F) = \Pr(E) + \Pr(F) - \Pr(E \cap F)$
6. Inclusion/exclusion principle.

Lec 1: Axiomatic Definition of Probability

Say we have an experiment where every one of N possible outcomes is likely. Let event E be some union of the outcomes. We introduced the following definition of probability:

$$\Pr(E) = \frac{\#\text{ of ways event } E \text{ can occur}}{N}$$

Lec 2: Combinatorics

Permutations and Combinations:

- Hypergeometric distribution (sampling from an urn without replacement)
- Birthday Problem

Lec 3: Conditional Probability + Multiplication Rule

The conditional probability of an event A, given the event B, is defined by

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

provided $\Pr(B) > 0$.

Lec 3: Conditional Probability + Multiplication Rule

The conditional probability of an event A, given the event B, is defined by

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

provided $\Pr(B) > 0$.

For any n events $E_1, \ldots, E_n$, the multiplication rule states

$$\Pr(E_1 \cap \ldots \cap E_n) = \Pr(E_n|E_1, \ldots, E_{n-1}) \times \ldots \times \Pr(E_2|E_1) \times \Pr(E_1)$$

For only two events:

$$\Pr(A \cap B) = \Pr(A|B) \times \Pr(B) = \Pr(B|A) \times \Pr(A)$$

Lec 4: Independence

We say that two events A and B are independent if

$$\Pr(A \cap B) = \Pr(A) \times \Pr(B)$$

Lec 4: Independence

We say that two events A and B are independent if

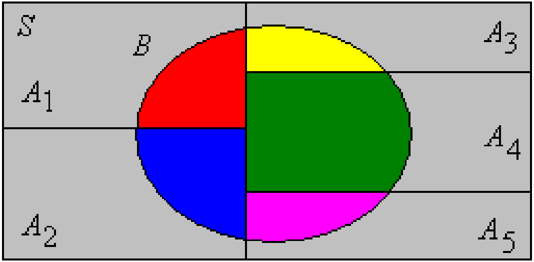
$$\Pr(A \cap B) = \Pr(A) \times \Pr(B)$$

Stated differently, say events A and B are independent, then

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A) \times \Pr(B)}{\Pr(B)} = \Pr(A)$$

Mutual independence: independence holds for any subset of events.

Lec 5: Law of Total Probability



$$
\begin{aligned}
\mathrm{Pr}(B) &= \mathrm{Pr}(B \cap A_1) + \mathrm{Pr}(B \cap A_2) + \mathrm{Pr}(B \cap A_3) + \\
&\quad \mathrm{Pr}(B \cap A_4) + \mathrm{Pr}(B \cap A_5) \\
&= \mathrm{Pr}(B|A_1) \cdot \mathrm{Pr}(A_1) + \mathrm{Pr}(B|A_2)\dot{\mathrm{P}}\mathrm{r}(A_2) + \mathrm{Pr}(B|A_3) \cdot \mathrm{Pr}(A_3) + \\
&\quad \mathrm{Pr}(B|A_4) \cdot \mathrm{Pr}(A_4) + \mathrm{Pr}(B|A_5) \cdot \mathrm{Pr}(A_5)
\end{aligned}
$$

Lec 5: Bayes Theorem (Simplified Form)

Let A and B be events such that $\Pr(A) > 0$, then Bayes theorem states:

$$\Pr(B|A) = \frac{\Pr(A|B) \cdot \Pr(B)}{\Pr(A)}$$

Lec 5: Bayes Theorem (Expanded Form with LOTP)

Now say we have a collection of events $B_1, \dots, B_k$ that form a partition of the sample space S, then for each $j = 1, \dots, k$

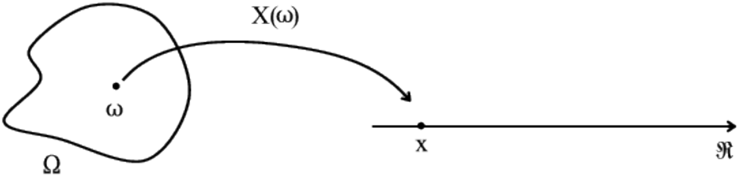
$$
\begin{aligned}
\Pr(B_j|A) &= \frac{\Pr(A|B_j) \cdot \Pr(B_j)}{\Pr(A)} \\
&= \frac{\Pr(A|B_j) \cdot \Pr(B_j)}{\sum_{i=1}^{k} \Pr(A|B_i) \cdot \Pr(B_i)}
\end{aligned}
$$

i.e. Apply the law of total probability to the denominator to normalize the distribution to sum to 1.

Lec 6: Random Variable

Definition

A function $X(\cdot)$ that maps the sample space S to the real line in such a way that, that for every $\omega \in S$, $X(\omega)$ is a real number, is called a random variable (RV for short).



Lec 6: Cumulative Distribution Function

Definition

We characterize the (random) behavior of a RV via its distribution function (AKA cumulative distribution function). The CDF of a random variable X is a function $F : \mathbb{R} \longrightarrow [0, 1]$ given by $F(x) = Pr(X \leq x)$

A function $F(x)$ is a CDF for some random variable X if and only if it satisfies the following properties

- $\lim_{x \to -\infty} F(X) = 0$
- $\lim_{x \to \infty} F(X) = 1$
- $\lim_{h \to 0^+} F(x + h) = F(x)$ (right continuous)
- $a < b$ also implies $F(a) \leq F(b)$

Lec 6: Cumulative Distribution Function

Lemmas: Let $F(x)$ be the CDF of X, then

- $\Pr(X > x) = 1 - F(x)$
- $\Pr(x < X \leq y) = F(y) - F(x)$
- $\Pr(X = x) = F(x) - \lim_{y \to x-} F(y)$

The quantile function is the inverse CDF function.

Lec 7: Discrete Random Variables

Definition

If the set of all possible values of a random variable X is a countable set, $x_1, x_2, \ldots$, then X is called a discrete random variable. The function

$$f(x) = Pr(X = x) \text{ for } x = x_1, x_2, \ldots$$

that assigns the probability to each possible value x will be called the probability density function (AKA probability mass function).

Lec 7: Properties

Theorem

A function $f(x)$ is a discrete PDF IFF it satisfies both of the following properties for at most a countably infinite set of real $x_1, x_2, \ldots$

1. $f(x_i) \geq 0$ *for all x_i: non-negative prob*
2. $\sum_{x_i} f(x_i) = 1$*: sums to one*

Lec 7 & 8: Discrete Distributions

- Bernoulli(p)
- Binomial(n, p)
- Geometric(p)
- Discrete Uniform
- Negative Binomial(r, p): both focusing on (random) number of trials and number of failures
- Poisson(λ): special case of Binomial as $n \longrightarrow \infty$ and $p \longrightarrow 0$

Lec 9: Expectation

If X is a discrete random variable with PDF $f(x)$, then the expected value of X is defined by

$$\mathbb{E}[X] = \mu = \sum_{x} x \cdot f(x)$$

Lec 9: Expectation

If X is a discrete random variable with PDF $f(x)$, then the expected value of X is defined by

$$\mathbb{E}[X] = \mu = \sum_x x \cdot f(x)$$

Theorem

(Linearity) If X is a (discrete) random variable with PDF $f(x)$, a and b are constants, and $g(x)$ and $h(x)$ are real-valued functions whose domains include the possible values of X, then

$$\begin{aligned}\mathbb{E}[a \cdot g(X) + b \cdot h(X)] &= a \cdot \mathbb{E}[g(X)] + b \cdot \mathbb{E}[h(X)] \\ &= a \cdot \sum_x g(x) \cdot f(x) + b \cdot \sum_x h(x) \cdot f(x)\end{aligned}$$

Lec 9: Variance and Standard Deviation

The variance σ^2 AKA Var(X) of a distribution is

$$\begin{aligned} \mathrm{Var}(X) &= \mathbb{E}\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2 \cdot f(x) \\ &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[X^2] - \mu^2 \end{aligned}$$

If X is a random variable and a and b are constants then

$$Var(aX+b) = a^2 Var(X)$$

Lec 10: Continuous Random Variable

A random variable X is called a continuous random variable if there is a function $f(x)$ called the probability density function (PDF) of X, such that the CDF can be represented as

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

By the Fundamental Theorem of Calculus:

$$f(x) = \frac{d}{dx}F(x)$$

Also

$$\Pr(a \leq X \leq b) = \int_{a}^{b} f(t)dt = F(b) - F(a)$$

Lec 10: Probability Density Function & Expectation

Theorem: A function $f(x)$ is a PDF for some continuous random variable X IFF if it satisfies the properties:

1. $f(x) \geq 0$ for all real x
2. $$\int_{-\infty}^{\infty} f(x)dx = 1$$

Lec 10: Probability Density Function & Expectation

Theorem: A function $f(x)$ is a PDF for some continuous random variable X IFF if it satisfies the properties:

1. $f(x) \geq 0$ for all real x
2. $$\int_{-\infty}^{\infty} f(x)dx = 1$$

If X is a continuous random variable with PDF $f(x)$ then the expected value of X is defined by

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)dx$$

Lec 11 & 12: Continuous Distributions

Simpler as

- Continuous Uniform(a, b)
- Gamma (know Gamma function)
- Exponential
- Normal(μ, σ^2)
- χ_p^2 with p degrees of freedom (A Normal($0, 1$) squared is χ_1^2)