Lecture 5.2: Midterm Review

2013/10/02

Lec 1.2: Experiments, Sample space, Events

- Experiments
- Sample space
- Events

Carefully label events of interest.

Lec 1.2: Kolmogorov Axioms

Let S be a sample space and let E be an event in S, then

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

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

Inclusion/exclusion principle.

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

Lec 1.2: Useful Consequences of the Kolmogorov Axioms

From

1. $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$
2. $\mathbb{P}(\emptyset) = 0$
3. If $E \subset F$, then $\mathbb{P}(E) \leq \mathbb{P}(F)$
4. $\mathbb{P}(E \cap F^c) = \mathbb{P}(E) - \mathbb{P}(E \cap F)$
5. For any two events E and F (not necessarily disjoint), we have $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$

Be prepared to make similar proofs using the Kolmogorov Axioms

Lec 1.2: Different Notion 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:

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

Lec 1.3: Permutations and Combinations

No questions directly on this stuff, but need to know it in so much as to be able to answer questions like the birthday problem.

Lec 2.1: Conditional Probability + Multiplication Rule

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

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

provided $\mathbb{P}(B) > 0$.

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

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

For only two events:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \times \mathbb{P}(B) = \mathbb{P}(B|A) \times \mathbb{P}(A)$$

Lec 2.2: Independence

We say that two events A and B are independent if

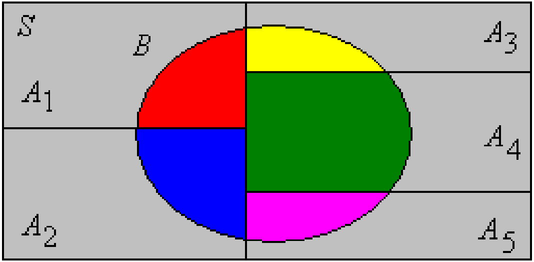
$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$$

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

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

Also, mutual independence: for any subset of events.

Lec 2.3: Law of Total Probability



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

Lec 3.1: Bayes Theorem (Simplified Form)

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

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

Lec 3.1: Bayes Theorem (Expanded Form)

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

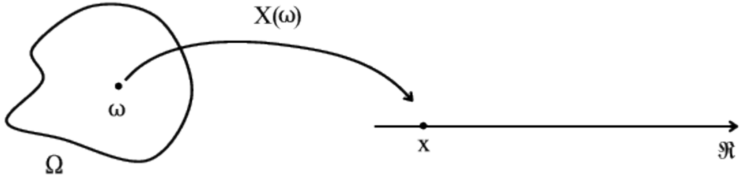
$$
\begin{aligned}
\mathbb{P}(B_j|A) &= \frac{\mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j)}{\mathbb{P}(A)} \\
&= \frac{\mathbb{P}(A|B_j) \cdot \mathbb{P}(B_j)}{\sum_{i=1}^{k} \mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)}
\end{aligned}
$$

i.e. Apply the law of total probability to the denominator

Lec 3.2: 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 3.2: Cumulative Distribution Function CDF

Definition

The distribution function (AKA cumulative distribution function) of a random variable X is a function $F : \mathbb{R} \longrightarrow [0,1]$ given by $F(x) = \mathbb{P}(X \leq x)$

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 3.2: Cumulative Distribution Function CDF

Let $F(\cdot)$ be the distribution function of X. Then

- $\mathbb{P}(X > x) = 1 - F(x)$
- $\mathbb{P}(x < X \leq y) = F(y) - F(x)$
- $\mathbb{P}(X = x) = F(x) - \lim_{y \uparrow x} \mathrm{F(y)}$

Lec 3.3: 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

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

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

Lec 3.3: 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 3.3: Discrete Distributions

- Bernoulli(p)
- Binomial(n, p)
- Geometric(p)
- Negative Binomial(r, p)
- Discrete Uniform
- Poisson(λ)

Lec 4.1: Expected Value + Linearity

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

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

Theorem

If X is a (discrete) random variable with PMF $f(x)$, 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 4.1: Variance

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 5.1: Probability Density Function

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

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

Lec 5.1: 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

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

Next Time... After Midterm

Introduce several continuous distributions

- Beta
- Exponential
- Uniform
- Gamma