Lecture 2

Monday, January 11, 2016
2:53 AM

Random variables and expectations

Underlying probability space consists of
1) sample space Ω
2) set of allowed events $\mathcal{F} \subseteq 2^{\Omega}$
3) A probability measure Pr

Ω consists of elementary events ω, all possible outcomes of an experiment.

- If experiment is tossing coin 2 times,
  $\Omega = \{HH, HT, TH, TT\}$
- If experiment is tossing coin until first head,
  $\Omega = \{H, TH, TTH, \dots\}$ (infinite)

An event e is a subset of Ω. Let $\mathcal{F} \subseteq 2^{\Omega}$ be the set of relevant events. ($\mathcal{F}$ has to satisfy the conditions of a σ- algebra but we ignore this here.)

A probability measure $Pr : \mathcal{F} \to \mathbb{R}$ satisfies:

(1) For any $e \in \mathcal{F}$, $0 \leq Pr(e) \leq 1$

(2) $Pr(\Omega) = 1$

(3) For any countable sequence of disjoint events $E_1, E_2, \dots$

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

We look at discrete probability spaces where Ω is countable. Here, we take $\mathcal{F} = 2^{\Omega}$.

Here's a consequence of the definitions.

Claim For any two events E_1 and E_2,

$$Pr[E_1 \cup E_2] = Pr[E_1] + Pr[E_2] - Pr[E_1 \cap E_2]$$

Pf:

$$Pr[E_1] = Pr[E_1 - E_1 \cap E_2] + Pr[E_1 \cap E_2]$$

$$Pr[E_2] = Pr[E_2 - E_1 \cap E_2] + Pr[E_1 \cap E_2]$$

$$Pr[E_1 \cup E_2] = Pr[E_1 - E_1 \cap E_2] + Pr[E_2 - E_1 \cap E_2] + Pr[E_1 \cap E_2]$$

Done. ☒

~~Corollary~~ $\boxed{\text{Union bound}}$: For any sequence of events:

$$\Pr\left[\bigcup E_i\right] \leq \sum \Pr[E_i]$$

Example: Freivald's matrix multiplication algorithm

Given $n \times n$ matrices A, B, C, does $AB = C$?

Idea: Choose random $r \in \{0,1\}^n$, check if $A(Br) = Cr$.

Claim: If $AB \neq C$, then $\Pr[\text{test says YES}] \leq \frac{1}{2}$.

Pf: Let $D = AB - C$. Suppose $D_{11} \neq 0$.

$$\Pr[Dr = 0] \leq \Pr\left[\sum_{j=1}^{n} D_{1j} r_j = 0\right]$$

$$= \Pr\left[r_1 = -\frac{\sum_{j=2}^{n} D_{1j} r_j}{D_{11}}\right] \leq \frac{1}{2}$$

Example: Ramsey number lower bound

Let $R(k,\ell)$ be smallest n s.t. for any graph G on n vertices, either G contains a clique of size k or independent set of size ℓ. Ramsey showed $R(k,\ell)$ finite.

Claim (Erdos '47): If $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$, then $R(k,k) > n$. In particular, $R(k,k) > \lfloor 2^{k/2} \rfloor$

Pf Take a random graph on n vertices, each edge present with prob. $1/2$.

For any fixed subset of vertices S of size k,

$$\Pr[S \text{ is clique or ind set}] = \frac{2}{2^{\binom{k}{2}}} = 2^{1-\binom{k}{2}}$$

By union bound,

$\Pr[\exists\, S$ of size k s.t. S is clique or ind set$]$

$$\le \binom{n}{k} \cdot 2^{1-\binom{k}{2}}$$

Since this < 1, There must exist G s.t. $\forall S$ of size k, S induces neither clique nor ind set.

A random variable X is a function $X : \Omega \to \mathbb{R}$.
For a discrete prob space, the range of X is countable.

Def: $$\mathbb{E}[X] = \sum_{i \in \text{range}(X)} i \cdot \Pr[X = i]$$

Expectation can be unbounded: 2^i with prob. 2^{-i}.

Claim: $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

Linearity of expectations

Pf: ...

Claim: If $\mathbb{E}[X] < k$, there must exist $\omega \in \Omega$ s.t. $X(\omega) < k$.

Example: A dominating set of a graph is a subset $S \subseteq V(G)$ s.t. $\forall v \in V(G)$, $v \in S$ or $(u,v) \in E(G)$ for some $u \in S$.

Claim: If G is a graph of size n with min degree $\ge \delta$, then G has a dominating set of size $\le \frac{1+\ln(1+\delta)}{1+\delta} n$

Pf: Choose each $v \in V$ with prob p and put it in Λ.

$\mathbb{E}[X] = np$. (Digression)

Take a vertex $u \in V$.

$\Pr[u \text{ not covered}] \leq (1-p)^{\delta}$

Let Y = # of uncovered vertices.

$\mathbb{E}[Y] \leq n(1-p)^{\delta+1}$

$$\mathbb{E}[X+Y] \leq n\left(p + (1-p)^{\delta+1}\right)$$
$$\leq n\left(p + e^{-p(\delta+1)}\right)$$
$$\leq n \cdot \frac{1+\ln(\delta+1)}{1+\delta}$$

by taking $p = \frac{\ln(1+\delta)}{1+\delta}$

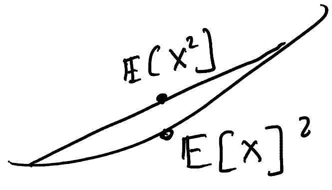
So, there exists dom set of this size!!

Jensen's Inequality

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

Note: First def immediately implies $\mathrm{Var}[X] \geq 0$

so, $\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$.



Generally true for any convex function (curve lies below line)

Fact: If f is twice differentiable, f is convex

Fact: If f is twice differentiable, f is convex $\iff f''(x) \geq 0 \ \forall x$.

Jensen's Inequality: $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$ if f is convex

Pf: Expand f using Taylor expansion around $\mu = \mathbb{E}[X]$

$$f(x) = f(\mu) + f'(\mu)\cdot(x-\mu) + \underbrace{f''(c)\cdot\frac{(x-\mu)^2}{2}}_{\geq 0}$$

$$\Rightarrow \mathbb{E}[f(x)] \geq f(\mu).$$

Conditional probabilities and independence

Def: Conditional probability that E occurs given F occurs is:

$$Pr[E \mid F] = \frac{Pr[E \cap F]}{Pr[F]} \text{ if } Pr[F] > 0.$$

Iow: $Pr[E \cap F] = Pr[E \mid F] \cdot Pr[F]$

Naturally, then, if X is a random var and E an event, then:

$$\mathbb{E}[X \mid E] \triangleq \sum_i i \cdot Pr[X = i \mid E].$$

Def: Events E and F are independent if

$$Pr[E \cap F] = Pr[E] \cdot Pr[F].$$

More generally, $E_1, \ldots, E_k$ are independent if

$$Pr[\cap E_i] = \prod Pr[E_i]$$

More generally, $E_1, \dots, E_k$ are independent if
$$\forall I \subseteq [k], \quad \Pr\left[\bigcap_{i \in I} E_i\right] = \prod_{i \in I} \Pr[E_i]$$

<u>Fact</u>: $\Pr[E \mid F] = \Pr[E]$ iff $E \perp F$.

<u>Example</u>: X_1 be outcome of one die roll and X_2 of another independent roll. $X = X_1 + X_2$

$$\mathbb{E}[X \mid X_1 = 2] = \sum_i i \cdot \Pr[X = i \mid X_1 = 2]$$
$$= \sum_{i=3}^{8} i \cdot \Pr[X_2 = i - 2]$$
$$= \sum_{i=3}^{8} i \cdot \frac{1}{6} = \frac{11}{2}$$

<u>Fact</u>: If $X \perp Y$, $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

<u>Pf</u>: Use $\mathbb{E}[X] = \sum_y \Pr[Y = y] \cdot \mathbb{E}[X \mid Y = y]$.

<u>Def</u>: For two random vars X, Y:

$\mathbb{E}[X \mid Y]$ is a function of y defined as $\mathbb{E}[X \mid Y = y]$.

<u>Claim</u>: $\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid Z]]$

<u>Pf</u>: $\mathbb{E}[\mathbb{E}[Y \mid Z]] = \sum \mathbb{E}[Y \mid Z = z] \cdot \Pr[Z = z]$
$$= \mathbb{E}[Y].$$

<u>Example (Branching process)</u>:

... program P once, and that call spawns

We call a program P once, and that call spawns new processes of P where number of new copies is a <u>binomial</u> random variable with params n and p. What is the expected # of total copies of P running?

Y_i = # of copies of P in generation i

$$Y_0 = 1$$

$$\mathbb{E}[Y_1] = np$$

$$\mathbb{E}[Y_i \mid Y_{i-1} = y_{i-1}] = \mathbb{E}\left[\sum_{k=1}^{y_{i-1}} Z_k \mid Y_{i-1} = y_{i-1}\right]$$

$$= y_{i-1}\, np$$

using $Z_k \perp Y_{i-1}$.

$$\mathbb{E}[Y_i \mid Y_{i-1}] = Y_{i-1} \cdot np$$

$$\Rightarrow \mathbb{E}[Y_i] = np\, \mathbb{E}[Y_{i-1}]$$

$$= (np)^i$$

$$\Rightarrow \mathbb{E}\left[\sum Y_i\right] = \frac{1}{1-np} \quad \text{if } np < 1$$

$$= \infty \quad \text{o.w.}$$