Expectation

1/20/22

L2 Ω: Sample space

Atomic event / elementary events

$\mathcal{E} \subseteq \Omega$

Event: Dice is even $= \{2,4,6\}$

$P: 2^{\Omega} \to [0,1]$:

- $P[\Omega] = 1$
- $\forall X, Y \subseteq \Omega$ $X \cap Y = \emptyset \Rightarrow P[X \cup Y] = P[X] + P[Y]$.

$P: F \to [0,1]$

$F \subseteq 2^{\Omega}$: σ-algebra.

Union bound

$\mathcal{E}_1, \mathcal{E}_2, \dots \mathcal{E}_m \subseteq \Omega$

$$P\left[\bigcup \mathcal{E}_i\right] \leq \sum_{i} P[\mathcal{E}_i]$$

Random variable RV

$f: \Omega \to T$

$X \in \mathbb{Z}$ $\quad i \in \mathbb{Z}$ $P[X=i]$

Expectation of a RV X is

$$\mu_X = E[X] = \sum_{x \in \text{all values that } X \text{ might get}} x \cdot P[X=x]$$ — average value of X

Conditional probability

X, Y: Two RV

$P[X=x \mid Y=y] =$ probability of X to be x given that Y is y.

$X=1 \iff$ the dice is even
$\quad = 0$ otherwise
$Y=1 \iff$ The dice ≤ 3
$\quad = 0$ otherwise

Indicator variables

$P[X=1 \mid Y=1] = P[d=2,4,6 \mid d=1,2,3] = \frac{1}{3}$

$P[X=1] = \frac{1}{2}$

Ω

$$\frac{P[X=x \wedge Y=y]}{P[Y=y]}$$

$$P[X=x \mid Y=y] = \frac{P[X=x \wedge Y=y]}{P[Y=y]}$$

Two random variables X and Y are independent if

$$P[X=x \mid Y=y] = P[X=x]$$

$$\iff P[X=x \wedge Y=y] = P[X=x]\,P[Y=y]$$

Expectation

Lemma $\forall X, Y$ RV we have

$E[X+Y] = E[X] + E[Y]$.

Linearity of expectation

E is a linear operator.

$E[X+Y] = E[X] + E[Y]$

$E[c \cdot X] = c\,E[X]$. (c: constant)

Lemma

$\forall X, Y$ RV that are independent

$E[XY] = E[X]E[Y]$.

Variance

$\mu = E[X]$

-1, 0, $+1$ with $\frac{1}{2}$, $\frac{1}{2}$ $\qquad$ -10, 0, $+10$ with $\frac{1}{2}$, $\frac{1}{2}$

$$V[X] = E[(X-\mu)^2]$$
$$= E[X^2 - 2X\mu + \mu^2]$$
$$= E[X^2] - 2E[X]\mu + \mu^2$$
$$= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2$$

variance = second moment.

Bernoulli distribution

Flip coin with prob p for heads: Get 1 otherwise 0.

$X \sim \text{Bernoulli}(p)$

$E[X] = 0 \cdot P[X=0] + 1 \cdot P[X=1] = 0 \cdot m + 1 \cdot p = p$.

$V[X] = pq$.

Binomial distribution

$b(k; n, p) = P[\text{getting } k \text{ heads}] = \binom{n}{k} p^k (1-p)^{n-k}$

(n: # flips; p: prob in the b trial)

$Y = X_1 + X_2 + \dots + X_n$, $\quad X_1, \dots, X_n$ are Bernoulli(p) independent

$E[Y] = np$

$V[Y] = V[\sum X_i] = \sum_i V[X_i] = npq \qquad q = (1-p)$.

(because X_is are independent.)

Geometric dist

$X \sim \text{Geom}(p)$ $\qquad X \in \{1, 2, 3, \dots\}$

$E[X] = \frac{1}{p}$

$V[X] = \frac{1-p}{p^2}$.

Approximating k-SAT

kSAT: a formula that is CNF

$\wedge(\sim)$

every clause is the $\vee$ or of exactly k literals, all literals in a clause use different variable.

F: Given kSAT formula
assume it uses n variable
and it has m clauses

max kSAT

Compute an assignment that satisfy the largest number of clauses.

Theorem 2SAT is NP-Hard.

PCP Theorem

Given a 3SAT formula with m clauses unless $P = NP$ one can not in polynomial time compute a satisfying assignment that satisfies more than $\frac{7}{8}m$ of the clauses.

$\left(\frac{7}{8} + \varepsilon\right)m$

F: 3SAT, m clauses

$C \equiv (x \vee y \vee z)$

Pick a random assignment $x \in \{0,1\}$ with equal prob independently

$P[C \text{ is satisfied}] \stackrel{?}{=} P[(x \vee y \vee z) = 1] = 1 - \frac{1}{8} = \frac{7}{8}$.

$P[(x \vee y \vee z) = 0] = P[x=0]\,P[y=0]\,P[z=0]$ (x, y, z all $=0$)

$= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$

$P[x \vee y \vee z = 0] = P[(x=0) \wedge (y=0) \wedge (z=0)]$
$= P[x=0]\,P[y=0]\,P[z=0]$.

$F = C_1 \wedge C_2 \wedge \dots \wedge C_m$

$P[C_i = 1 \text{ (its satisfied)}] = \frac{7}{8}$

$X_i = 1 \iff C_i$ is satisfied
$\quad = 0$ otherwise

X_i is an indicator variable.

$Y = \#$ of satisfied clauses $= \sum_{i=1}^{m} X_i$

$E[Y] =$ how many clauses are satisfied on average

$= E\left[\sum_{i=1}^{m} X_i\right] = \sum_{i=1}^{m} E[X_i] = \sum_{i=1}^{m} \frac{7}{8} = \frac{7}{8}m$.

(by linearity of expectations)

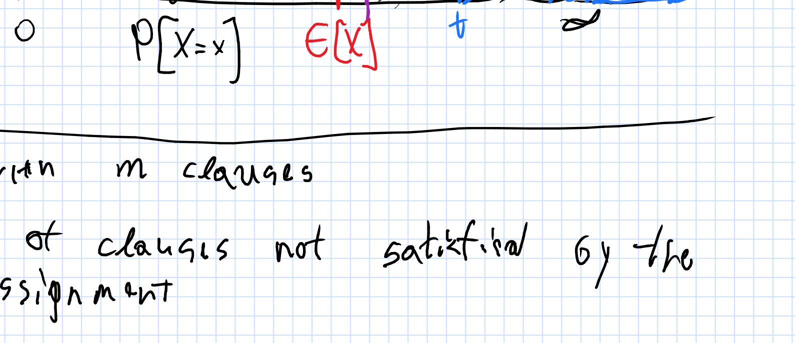
$E[X_i] = 0 \cdot P[X_i = 0] + 1 \cdot P[X_i = 1]$
$= P[X_i = 1] = \frac{7}{8}$

Markov inequality

X: RV with real values

$$P[X \geq t] \leq \frac{E[X]}{t}. \qquad t \geq E[X]$$

proof



kSAT F with m clauses

$Z = m - Y = \#$ of clauses not satisfied by the assignment

$E[Z] = m - \frac{7}{8}m = \frac{1}{8}m$

$\varepsilon \in (0,1)$ $\qquad \varepsilon = 0.1$

$$P\left[Z \geq (1+\varepsilon)\tfrac{1}{8}m\right] = P[Z \geq (1+\varepsilon)E[Z]] \leq \frac{E[Z]}{(1+\varepsilon)E[Z]}$$

$\frac{1}{8}(1+\varepsilon)$ $\qquad = \frac{1}{1+\varepsilon} \leq 1 - \frac{\varepsilon}{2}$

$\leq \left(1 - \frac{\varepsilon}{2}\right)^t \to 0$

$t = O\left(\frac{1}{\varepsilon}\right)$