Lecture 7 - 9/7

Review:

① Random variables $X : \Omega \to \mathbb{R}$

. Discrete RV

. Support, realization,

. Probability mass function

② Mean, Variance, Standard deviation

. $\mu_X = E[X] = \sum_x x\, p_X(x)$

. $Var(X) = E\left[(X-\mu_X)^2\right] = E\left[X^2\right] - \mu_X^2$

. $\sigma_X = \sqrt{Var(X)}$

. LOTUS and implications

③ Conditional probability:

. $P(B|A) = \begin{cases} \dfrac{P(AB)}{P(A)}, & \text{if } P(A) > 0 \\ \text{undefined}, & \text{if } P(A) = 0 \end{cases}$

. It is a probability measure, i.e., satisfies axioms.

④ Independent events:

. A and B are independent if $P(AB) = P(A)P(B)$

. A, B, C are pairwise independent if

$P(AB) = P(A)P(B)$, $P(AC) = P(A)P(C)$, $P(BC) = P(B)P(C)$

. A, B, C are pairwise independent if

$$P(AB) = P(A)P(B), \quad P(AC) = P(A)P(C), \quad P(BC) = P(B)P(C)$$

. A, B, C are independent if they are pairwise independent &

$$P(ABC) = P(A)P(B)P(C)$$

. Consider picking an element from $\{1,2,3,\dots,9\}$ at random.

$$A = \{1,2,3\}, \quad B = \{3,4,5\}, \quad C = \{5,2,6\}$$

A, B, C are pairwise independent; however $P(ABC) = 0$

Today:

1. Independent random variables
2. Important remarks
3. Bernoulli distribution
4. Binomial distribution

---

① independent discrete random variables:

Def: We say X and Y are independent if $\{X \in A\}$ and $\{Y \in B\}$ are independent for any A and B. Recall $\{X \in A\} = \{\omega \in \Omega : X(\omega) \in A\}$

. X and Y are independent $\Rightarrow P(X=i, Y=j) = p_X(i)\, p_Y(j)$ for all i and j

. Conversely, if $P(X=i, Y=j) = p_X(i)\, p_Y(j)$ for all i and j then:

$$P(X \in A, Y \in B) = \sum_{(i,j):\, i \in A,\, j \in B} P(X=i, Y=j)$$

$$P(X\in A, Y\in B) = \sum_{(i,j):\, i\in A,\, j\in B} P(X=i, Y=j)$$

$$= \sum_{(i,j):\, i\in A,\, j\in B} P_X(i)P_Y(j) = P(X\in A)P(Y\in B)$$

. Note that we are talking about discrete random variables.

---

② Important remarks:

2.1 Mutually exclusive vs independent

. A & B are mutually exclusive if $A\cap B = \emptyset$

. A & B are independent if $P(AB) = P(A)P(B)$

2.2 Intersection vs Conditional Probability

. $P(AB)$: probability of both A & B happening

. $P(B|A)$: we already know A has happened, what is the probability of B happening?

All we are interested in is statistical properties. Hence, we only care about distributions, i.e., pmf for discrete random variables. Some distributions that are frequent have names.

① Binomial distribution

. Toss a coin, outcome is T or H. $X(T)=1$, $X(H)=0$.

. Digital transmission, outcome is bit 1 or bit 0. $X(\text{bit } 1)=1$, $X(\text{bit } 0)=0$.

Def. We say random variable X has Bernoulli distribution with parameter $0\le p\le 1$

Def. We say random variable X has Bernoulli distribution with parameter $0 \leq p \leq 1$ if: $p_X(1) = P(X=1) = p$, $p_X(0) = P(X=0) = 1-p$

- $E[X^2] = E[X] = p$, $Var(X) = p(1-p)$

② Binomial distributions

. Toss a coin n times, X = number of heads.

. A bit stream of length n, X = number of ones.

Def. Consider n independent Bernoulli trials, each resulting in one with probability $0 \leq p \leq 1$.

We have

$$P(k \text{ out of } n \text{ trials is one}) = \underbrace{\binom{n}{k}}_{\text{picking location of } k \text{ trials}} \times \underbrace{p^k}_{\text{prob. of one in } k \text{ trials}} \times \underbrace{(1-p)^{n-k}}_{\text{prob. of zero in } n-k \text{ trials}}$$

We say a random variable X has Binomial distribution with parameters (n,p) if its PMF is given by:

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } k \in \{0,1,2,\dots,n\}$$

- We need to check P_X satisfies axioms of probability. In particular, we need to check $\sum_{k=0}^{n} p_X(k) = 1$, as other axioms is trivial.

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (1-p)^n \sum_{k=0}^{n} \left(\frac{p}{1-p}\right)^k$$

Recall the binomial expansion: $(a+b)^n = \sum^{n} \binom{n}{k} a^k b^{n-k}$.

Recall the binomial expansion: $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$.

Replacing $a=1$ and $b = \frac{p}{1-p}$, we get

$$\left(1+\frac{p}{1-p}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{p}{1-p}\right)^k$$

So we have:

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (1-p)^n \sum_{k=0}^{n} \binom{n}{k} \left(\frac{p}{1-p}\right)^k$$

$$= (1-p)^n \cdot \left(1+\frac{p}{1-p}\right)^p = 1$$

- Next we derive statistical properties of X:

(i) Mean of X: we present 3 ways to derive mean of a binomial random variable

Method 1: As we mentioned, a binomial random variable is number of successes in n independent trials. Let X_i denote the result of ith trial; $X_i = 1$ if ith trial was success and $X_i = 0$ otherwise. We have

$$X = X_1 + X_2 + \cdots + X_n$$

By linearity of expectation, we have

$$E[X] = E[X_1] + E[X_2] + \cdots + E[X_n] = np$$

Method 2: we can directly calculate the value. The idea is to

Method 2: we can directly calculate the value. The idea is to write everything in terms of binomial expansion.

$$E[X] = \sum_{k=0}^{n} k\, p_X(k)$$

$$= \sum_{k=0}^{n} k \cdot \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!}\, p^k (1-p)^{n-k} \rightsquigarrow \text{notice change in summand}$$

$$= \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k (1-p)^{n-k} = np \underbrace{\sum_{\ell=1}^{n-1} \binom{n-1}{\ell} p^{\ell} (1-p)^{n-1-\ell}}_{=1} = np$$

Method 3: Given the binomial expansion $(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$, taking derivative with respect to x, we get:

$$n(1+x)^{n-1} = \sum_{k=1}^{n} \binom{n}{k} k\, x^{k-1}$$

The idea is to rewrite $E[X]$ as above:

$$E[X] = \sum_{k=0}^{n} k\, p_X(k) \rightsquigarrow \text{we can drop } k=0 \text{ from range}$$

$$= \sum_{k=1}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

$$= (1-p)^n \cdot \frac{p}{1-p} \cdot \sum_{k=1}^{n} k \binom{n}{k} \left(\frac{p}{1-p}\right)^{k-1}$$

$$= p(1-p)^{n-1} \cdot n\left(1+\frac{p}{1-p}\right)^{n-1} = np$$

$$= p(1-p)^{n-1} \cdot n\left(1+\frac{p}{1-p}\right)^{n-1} = np$$

(ii) Variance of X. we can replicate all three methods for variance as well. However, this requires a bit more knowledge on independence. We only focus on last two methods. We only need to calculate $E[X^2]$.

Method 1: direct calculation

$$E[X^2] = \sum_{k=0}^{n} k^2 p_X(k)$$

$$= \sum_{k=1}^{n} k^2 \frac{n!}{k!\,(n-k)!} \cdot p^k \cdot (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} k \frac{n!}{(k-1)!\,(n-k)!} p^k (1-p)^{n-k}$$

$$= np \sum_{\ell=0}^{n-1} (\ell+1) \binom{n-1}{\ell} p^\ell (1-p)^{n-1-\ell}$$

$$= np \Big(\underbrace{\sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^\ell (1-p)^{n-1-\ell}}_{=1} + \underbrace{\sum_{\ell=0}^{n-1} \ell \binom{n-1}{\ell} p^\ell (1-p)^{n-1-\ell}}_{\text{mean of Binomial random variable with parameters } (n-1,\,p)} \Big)$$

$$= np\,(1+(n-1)p) = np + n^2p^2 - np^2$$

Method 2: taking another derivative from binomial expansion, we get

$$n(n-1)(1+x)^{n-2} = \sum_{k=2}^{n} k(k-1) \binom{n}{k} x^{k-2}$$

$$= \sum_{k=2}^{n} k^2 \binom{n}{k} x^{k-2} - \sum_{k=2}^{n} k \binom{n}{k} x^{k-2}$$

$$= \sum_{K=2}^{n} K^2 \binom{n}{k} x^{K-2} - \sum_{K=2} K \binom{n}{k} x^{K-2}$$

Multiplying both sides with x,

$$n(n-1)x(1+x)^{n-2} = \sum_{K=2}^{n} K^2 \binom{n}{k} x^{K-1} - \sum_{K=2}^{n} K \binom{n}{k} x^{K-1}$$

$$= \sum_{K=1}^{n} K^2 \binom{n}{k} x^{K-1} - \sum_{K=1}^{n} K \binom{n}{k} x^{K-1}$$

$$= \sum_{K=1}^{n} K^2 \binom{n}{k} x^{K-1} - n(1+x)^{n-1}$$

Hence, we have: $\displaystyle\sum_{K=1}^{n} K^2 \binom{n}{k} x^k = n(1+x)^{n-2}\left((n-1)x + (1+x)\right)x$

$$E[X^2] = \sum_{K=0}^{n} K^2 \, P_X(k)$$

$$= \sum_{K=0}^{n} K^2 \binom{n}{k} p^K (1-p)^{n-K}$$

$$= (1-p)^n \sum_{K=1}^{n} K^2 \binom{n}{k} \left(\frac{p}{1-p}\right)^K$$

$$= (1-p)^n \cdot n \cdot \left(1 + \frac{p}{1-p}\right)^{n-2} \left((n-1) \cdot \frac{p}{1-p} + \left(1 + \frac{p}{1-p}\right)\right)\left(\frac{p}{1-p}\right)$$

$$= n(1-p)\left(\frac{(n-1)p + 1}{1-p}\right)p = np + n^2p^2 - np$$

$$\rightsquigarrow Var(X) = E[X^2] - (E[X])^2 = np(1-p)$$

(iii) Mode of PMF: Mode of PMF is the most probable realization of random variable.

Notice that, for Binomial distribution, for any $1 \leq K \leq n$, we have:

Notice that, for Binomial distribution, for any $1 \leq k \leq n$, we have:

$$\frac{P_X(k)}{P_X(k-1)} = \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\binom{n}{k-1} p^{k-1} (1-p)^{n-k+1}} = \frac{\frac{n!}{(n-k)!\, k!} \cdot p}{\frac{n!}{(k-1)!\,(n-k+1)!} \cdot (1-p)} = \frac{(n-k+1)p}{k(1-p)}$$

Hence, if $P_X(k) \geq P_X(k-1)$, we have $(n-k+1)p \geq k(1-p)$ which implies $k \leq (n+1)p$. Since k is an integer, this implies $k \leq \lfloor (n+1)p \rfloor$.

$$\Rightarrow \text{mode of pmf of Binomial r.v. } X = \underset{k \in \{0,1,\dots,n\}}{\operatorname{argmax}} P_X(k) = \lfloor (n+1)p \rfloor$$

(iv) Median of Binomial random variable:

Median of a Binomial random variable is the smallest $\hat{k}$ for which

$$\sum_{k=0}^{\hat{k}} P_X(k) \geq 0.5$$

intuitively speaking, median is the value for which the probability that realization is larger than $\hat{k}$ is almost equal to the probability being smaller than or equal to $\hat{k}$. It can be shown that for Binomial random variable:

$$|\hat{k} - \mu_X| \leq \max\{p, 1-p\}$$

where $\mu_X = np$ is the mean of random variable X.