Lec 9&10

Independence of r.v.

R.V. X, Y are said to be independent if:

$$P(X\le x, Y\le y) = P(X\le x)\,P(Y\le y) \text{ for } \forall x, y \in \mathbb{R}$$

进阶: R.V. $X_1, \cdots, X_n$ are independent if

$$P(X_1\le x_1, \cdots, X_n\le x_n) = P(X_1\le x_1)\cdots P(X_n\le x_n)$$

for $\forall x_1, \cdots, x_n \in \mathbb{R}$

No.

Date

I.I.D. => independent and identically distributed

关于是否独立 & 是否相同分布，则X、Y关系就有四种

Binomial Distribution 二项分布

Theorem: if $X \sim Bin(n, p)$，则可视为 n 个独立Bernoulli实验，
则 $X = X_1 + \cdots + X_n$，X_i 是 i.i.d. $Bern(p)$

Theorem: If $X \sim Bin(n,p)$，$Y \sim Bin(m, p)$，X与Y独立，
则 $X+Y \sim Bin(m+n, p)$ *: 因为X、Y独立

Proof: $P(X+Y=k) = \sum_{j=0}^{k} P(X+Y=k \mid X=j)\, P(X=j)$

$$= \sum_{j=0}^{k} P(Y=k-j)\, P(X=j) = \sum_{j=0}^{k} \binom{m}{k-j} p^{k-j} (1-p)^{m-k+j} \cdot \binom{n}{j} p^{j} (1-p)^{n-j}$$

$$= \sum_{j=0}^{k} \binom{m}{k-j}\binom{n}{j} p^{k} (1-p)^{m+n-k}$$

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

* story proof: m男n女，共取k人，可视为男取k-j人，女取j人，j∈[0,k]

Conditional Independence of R.V.s

Def: R.V. X, Y are conditionally independent given an r.v. Z if for all $x, y \in \mathbb{R}$ and all z in the support of Z,

$$P(X \le x, Y \le y \mid Z=z) = P(X \le x \mid Z=z)\, P(Y \le y \mid Z=z)$$

Conditional PMF

Def: For any discrete r.v.s X & Z, function $P(X=x \mid Z=z)$, when considered as a function of x for fixed z, is called the conditional PMF of X given $Z=z$ (z是constant, x在变)

Binomial & Hypergeometric

Theorem: If $X \sim Bin(n,p)$, $Y \sim Bin(m,p)$, X. Y 间独立,
则 $X+Y=r$ 的条件分布为 $HGeom(n,m,r)$

Theorem: If $X \sim HGeom(w,b,n)$ and $N = w+b \to \infty$ s.t.
$p = w/(w+b)$ remains fixed, 则 X 的 PMF 收敛至 $Bin(n,p)$ PMF

Proof1: $X+Y \sim Bin(n+m, p)$ * 为什么看这个PMF? Story: HGeom: n+m中取r, r中有?个X 的分布 ⇔ $X+Y=r$ 下, $X=x$ 的条件PMF

$$P(X=x \mid X+Y=r)^{*} = \frac{P(X=x,\ X+Y=r)}{P(X+Y=r)}$$

$$= \frac{P(X=x,\ Y=r-x)}{P(X+Y=r)} = \frac{P(X=x)\,P(Y=r-x)}{P(X+Y=r)}$$

$$= \frac{\binom{n}{x} \cancel{p^{x}} \cancel{(1-p)^{n-x}} \binom{m}{r-x} \cancel{p^{r-x}} \cancel{(1-p)^{m-r+x}}}{\binom{m+n}{r} \cancel{p^{r}} \cancel{(1-p)^{m+n-r}}} = \frac{\binom{n}{x}\binom{m}{r-x}}{\binom{m+n}{r}} = HGeom(n,m,r)$$

Lec 9 Expectation

Def: The expected value of a discrete r.v. X:

$$E(X) = \sum_{j=1}^{\infty} x_j P(X=x_j)$$

Theorem: If X & Y 是有相同分布的离散 r.v.s, 则 $E(X)=E(Y)$

Theorem: Linearity: $E(X+Y) = E(X)+E(Y)$
$E(cX) = cE(X)$

Expectation via Survival Function:
Let X be a nonnegative integer-valued r.v. Let F be the CDF of X, and $G(x) = 1-F(x) = P(X>x)$.
The G is called the survival function of X. 则:

$$E(X) = \sum_{n=0}^{\infty} G(n)$$

No.

Date

这条性质很灵性：$x=1$ 只覆盖一次，$x=2$ 二次，…，$x=i$，i 次

Law of the unconscious statistician (LOTUS)

$$E(g(X))=\sum_x g(x)\,P(X=x)$$

Variance & Standard Deviation

$$Var(X)=E(X^2)-(E(X))^2$$

$$SD(X)=\sqrt{Var(X)}$$

Geometric Distribution

考虑一系列 Bernoulli 试验，概率为 $p\in(0,1)$，试验一直做直至一次成功。X是在成功前失败的次数，则认为 $X\sim Geom(p)$

PMF: $P(X=k)=q^k p$

Property: Memory less

$$P(X\geq n+k \mid X\geq k)=P(X\geq n)$$

vice versa：若离散 r.v. X满足 $P(X\geq n+k \mid X\geq k)=P(X\geq n)$

则 $X\sim Geom(p)$

Proof: ① $k=0$ ✓ ② $k\geq 1$: $P(X\geq n+k \mid X\geq k)$

$$=\frac{P(X\geq n+k,\,X\geq k)}{P(X\geq k)}=\frac{P(X\geq n+k)}{P(X\geq k)}=P(X\geq n)$$

令 $G(n)=P(X\geq n)$，$G(0)=1$，则欲证：$G(n+k)=G(n)G(k)$

$\because$ $G(n)=1-\sum_{i=0}^{n-1}P(X=i)=1-p(q^0+q^1+\cdots+q^{n-1})$

$$=1-p\cdot\frac{1-q^n}{1-q}=q^n$$

$\therefore$ $q^{n+k}=q^n\cdot q^k$ 显然成立

首次成功分布: Bernoulli试验, p是成功概率, Y是至首次成功时做的总试验次数。则 $Y \sim FS(p)$.

Negative Binomial Distribution: Bernoulli实验, X是在r^{th}成功前失败的次数, 则说X是$r.p$的负二项分布:

$$X \sim NBin(r,p)$$

PMF: $P(X=n) = \binom{n+r-1}{r-1} p^r q^n$, $q=1-p$

若$X \sim NBin(r,p)$. 则 $X = X_1 + \cdots + X_r$, 其中X_i是i.i.d $Geom(p)$.

欲说明该定理, 可用 Story Proof

Expectation: $X \sim Geom(p)$ | $X \sim FS(p)$ | $X \sim NBin(r,p)$

Coupon Collector: n种toy, 你想集齐。你一次能收集一张
N为你收集的次数. 求$E(N)$ & $Var(N)$

Solution: 我们令N_i代表: 收集到第i个toy所额外购买的toy

如: $\underbrace{\bigcirc}_{N_1}\ \underbrace{\bigcirc\ \square}_{N_2}\ \underbrace{\square\ \square\ \bigcirc}_{N_3}\ \triangle$

把这个问题拆成了一个个小FS分布问题 (建模?)

对于N_i来说, 一次Bernoulli成功: $p = \frac{n-i+1}{n}$, 则$E(N_i) = \frac{n}{n-i+1}$

$\therefore E(N) = \sum_{i=1}^{n} E(N_i) = n\left(1 + \frac{1}{2} + \cdots + \frac{1}{n-i+1}\right)$

$n \to \infty$: $\sum_{i=1}^{\infty} \frac{1}{n-i+1} = \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n-i+1}\right) = \int_1^{+\infty} \frac{1}{n} dn = \ln n$

$\therefore n \to \infty$下, $E(N) \sim n \ln n \approx n(\ln 2 + 0.577)$

No.

Date

Properties of Indicator R.V: $I\{A\}=\begin{cases}1, & 若A发生\\ 0, & 若A不发生\end{cases}$

$(I_A)^k=I_A$ $\quad I_{A^c}=1-I_A$ $\quad I_{A\cap B}=I_A I_B$ $\quad I_{A\cup B}=I_A+I_B-I_A I_B$

Theorem: $P(A)=E(I_A)$，架起了Probability & Expectation 桥梁

但左右两者计算量不同！在多变量 $\cap\cup$ 下，利用I.r.v 的property与期望线性性，$E(I)$ 计算更方便！

Eg: 证: Booler's Inequality:

$$P\left(\bigcup_{i=1}^{n} A_i\right)\leq\sum_{i=1}^{n}P(A_i)$$

欲证: $I(A_1+\cdots+A_n)\leq I(A_1)+\cdots+I(A_n)$

① if LHS=0 ，✓

② if LHS =1，说明 $A_1,\cdots,A_n$ 间至少有一个 A_i，$I(A_i)=1$

$\therefore$ RHS≥1，RHS$\geq$LHS

Lec 11

Moments and Indicators:

设 events $A_1,\cdots,A_n$， indicators I_j, $j=1,\cdots,n$.

$X=\sum_{j=1}^{n}I_j$ 代表 events 发生数量

$\binom{X}{2}=\sum_{i<j}I_iI_j$ 代表 发生的不同事件对的数量

则 $E\left(\binom{X}{2}\right)=\sum_{i<j}P(A_i\cap A_j)$

$$E(X^2)=2\sum_{i<j}P(A_i\cap A_j)+E(X)$$

$$Var(X)=2\sum_{i<j}P(A_i\cap A_j)+E(X)-(E(X))^2$$

$\triangle$: $E\left[\binom{X}{k}\right]=\binom{n}{k}p^k$

Poisson Distribution:

PMF: $P(X=k)=\dfrac{e^{-\lambda}\lambda^k}{k!}$，$X\sim Pois(\lambda)$.

Campus

核心: $\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}=1$，有点无穷级数味道

Theorem: ① $E(X)=\lambda$

$$E(X)=\sum_{k=0}^{\infty}k\cdot\frac{e^{-\lambda}\lambda^k}{k!}=e^{-\lambda}\cdot\lambda\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}$$

$$=e^{-\lambda}\cdot\lambda\cdot e^{\lambda}=\lambda$$

② $Var(X)=\lambda$

Poisson 近似: Let $A_1,\cdots,A_n$ 是 $p_j=P(A_j)$，当 n 很大，p_j 很小，A_j 独立时，

$$X=\sum_{j=1}^{n}I(A_j)\text{ 近似于 } Pois(\lambda).$$

其中 $\lambda=\sum_{j=1}^{n}p_j$

Theorem: if $X\sim Pois(\lambda_1)$ $Y\sim Pois(\lambda_2)$. X与Y独立，则 $X+Y\sim Pois(\lambda_1+\lambda_2)$

且: $X+Y=n$ 的 X 条件分布为 $Bin(n,\lambda_1/(\lambda_1+\lambda_2))$

Theorem: If $X\sim Bin(n,p)$，并令 $n\to\infty$, $p\to 0$, s.t. $\lambda=np$ remains fixed，则 X 的 PMF 收敛至 $Pois(\lambda)$ PMF

Distance between Two Probability Distributions

Def: $\mu.\nu$ 两个分布间的 total variation distance:

$$d_{TV}(\mu,\nu)=\|\mu-\nu\|_{TV}$$

$$=\frac{1}{2}\sum_{x\in\Omega}|\mu(x)-\nu(x)|$$

The Law of Rare Events.

有独立 r.v. $Y_1,\cdots,Y_n$ s.t. $\forall\ 1\le m\le n$, $P(Y_m=1)=p_m$ & $P(Y_m=0)=1-p_m$, $S_n=\sum_{i=1}^{n}Y_i$. 设:

$$\sum_{m=1}^{n}p_m\to\lambda\in(0,\infty)\text{ as }n\to\infty$$

No.

Date

若: $\max_{1\le m\le n} p_m \to 0$ as $n\to\infty$

则 $d_{TV}(S_n, Poi(\lambda)) \to 0$ as $n\to\infty$

Probability Generating Function.

The PGF of r.v. X with PMF $p_k = P(X=k)$ is the generating function of PMF:

$$E(t^X) = \sum_{k=0}^{\infty} p_k t^k.$$