基础数理统计 2023 Spring

Discrete dist.	pmf	mean	variance	mgf/moment
Discrete Uniform(n)	$\frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\frac{e^t(1-e^{nt})}{n(1-e^t)} = \frac{1}{n}\sum e^{it}$
Bernoulli(p)	$p^x(1-p)^{1-x}$	p	$p(1-p)$	$(1-p)+pe^t$
Binomial(n,p)	$\binom{n}{x}p^x(1-p)^{n-x}$	np	$np(1-p)$	$((1-p)+pe^t)^n$
Geometric(p)	$(1-p)^{x-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pe^t}{1-(1-p)e^t}$
Poisson(λ)	$\frac{\lambda^x}{x!}e^{-\lambda}$	λ	λ	$e^{\lambda(e^t-1)}$
Beta-binomial (n,α,β)	$\binom{n}{x}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\cdot \frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(n+\alpha+\beta)}$	$\frac{n\alpha}{\alpha+\beta}$	$\frac{n\alpha\beta}{(\alpha+\beta)^2}$	If $X\|P$ is Binomial(n,P), and P is Beta(α,β), then X is Beta-binomial(n,α,β).

Continuous dist.	pdf	mean	variance	mgf/moment
Uniform(a,b)	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb}-e^{ta}}{t(b-a)}$
Exponential(θ)	$\frac{1}{\theta}e^{-\frac{x}{\theta}}$	θ	θ^2	$\frac{1}{(1-t\theta)^2}, t<\frac{1}{\theta}$
Exponential(λ)	$\lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{\lambda}{\lambda-t}, t<\lambda$
Normal(μ,σ^2)	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	$e^{\mu t+\frac{\sigma^2t^2}{2}}$
Gamma(α,β)	$\frac{1}{\Gamma(\alpha)\beta^\alpha}x^{\alpha-1}e^{-x/\beta}$	$\alpha\beta$	$\alpha\beta^2$	$\left(\frac{1}{1-\beta t}\right)^\alpha, t<\frac{1}{\beta}$
Beta(α,β)	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$1+\sum_{k=1}^{\infty}\left(\prod_{i=0}^{k-1}\frac{\alpha+i}{\alpha+\beta+i}\right)\frac{t^k}{k!}$
Cauchy	$1/\left(\pi(1+x^2)\right), x\in\mathbb{R}$	(∞)	n/a	n/a, $F(x)=\arctan(x)/\pi+1/2$
χ^2_p	$\frac{1}{2^{p/2}\Gamma(p/2)}x^{p/2-1}e^{-x/2}$	p	$2p$	$(1-2t)^{-p/2}, t<1/2$

1.概率

条件概率 $P(A|B)=P(AB)/P(B)$
全概率 $P(B)=\sum_i P(B|A_i)P(A_i)$
贝叶斯 $P(A_i|B)=\frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)}$
作业：蓝眼 1/4。对于三个孩子的家庭，如果至少有一个孩子是蓝眼睛 (A)，至少有两个蓝眼睛 (B) 的概率为 $P(B|A)=P(AB)/P(A)=P(B)/P(A)=10/37$.
作业：$p_{1\ldots5}=0,\frac{1}{4},\frac{1}{2},\frac{3}{4},1$ 代表第 i 枚硬币出现正面的概率。随机选取一枚硬币，投掷直到出现正面。求 $P\left(C_i \mid B_4\right)$, B_4 表示第 4 次首次出现正面。由贝叶斯, $P\left(C_i|B_4\right)=\frac{P(B_4|C_i)P(C_i)}{\sum_{i\in[5]}P(B_4|C_i)P(C_i)}=\frac{P(B_4|C_i)}{\sum_{i\in[5]}P(B_4|C_i)}$.

2.随机变量

随机变量 $X:\Omega\mapsto\mathbb{R}$ 对每个样本赋予实值。
CDF $F_X(x)=P\left(X\le x\right)$
PDF 1: $f_X(x)\ge 0$, 2: $\int_{-\infty}^{+\infty}f_X(x)\,\mathrm{d}x=1$, 3: $P\left(a<X<b\right)=\int_a^b f_X(x)\,\mathrm{d}x$.
$F^{-1}(q)=\inf\{x:F(x)>q\}, (\max x: f(x)\le q)$
Poisson 分布：平均 5 分钟 10 人到店，问等待第一个顾客两分钟以上的概率。两分钟实际到店人数 X 服从参数为 $\lambda=4$ 的 Poisson 分布，$P(T>2)=P(X=0)=e^{-4}$. 推广，等待时间 $F(t)=1-e^{-\lambda t}, f(t)=F'(t)=\lambda e^{-\lambda t}$，为指数分布，$t\in[0,\infty)$。
二维正态分布：$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\cdot\exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}+\frac{(y-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right]\right)$
边际分布：$f_X(x)=\int_{-\infty}^{+\infty}f(x,y)\,\mathrm{d}y$
独立定义 $P(X\in A, Y\in B)=P(X\in A)P(Y\in B)$, $f_{X,Y}(x,y)=f_X(x)f_Y(y)$, 独立 iff $\forall x,y, f(x,y)=r(x)g(y), f_X(x)=r(x)/\int_{-\infty}^{\infty}r(x)\,\mathrm{d}x$
条件分布 $f_{X|Y}(x|y)=\frac{P(X=x,Y=y)}{P(Y=y)}=\frac{f_{X,Y}(x,y)}{f_Y(y)}$
多项分布：$P(\ldots,X_k=n_k)=\frac{n!}{n_1!\ldots n_k!}p_1^{n_1}\ldots p_k^{n_k}$
多元正态分布：$f(x_1,\ldots,x_n)=\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$
随机变量的函数：对于随机变量 X 考虑函数形式 $r(x)$，计算 $Y=r(X)$ 的分布：对每个 y 求集合 $A_y=\{x: r(x)\le y\}$；求 CDF: $F_{r(X)}(y)=P\left(r(X)\le y\right)=P\left(X\in A_y\right)=\int_{A_y}f_X(x)\,\mathrm{d}x$；求 PDF: $f_Y(y)=F'_Y(y)$。当 r 单调时，r 的反函数为 $s=r^{-1}$，有 $f_Y(y)=f_X(s(y))\left|\frac{\mathrm{d}s(y)}{\mathrm{d}y}\right|$。
例：$f_X(x)=e^{-x}(x>0), F_X(x)=\int_0^x f_X(s)\,\mathrm{d}s=1-e^{-x}$。令 $Y=r(X)=\log X, A_y=\{x:x\le e^y\}$, $F_Y(y)=P\left(X\le e^y\right)=F_X(e^y)=1-e^{-e^y}$, $f_Y(y)=F'_Y(y)=e^y e^{-e^y}$。
作业：令 $X,Y\sim$ Uniform$(0,1)$ 且独立，求 $X-Y$ 的 PDF。令 $Z=X-Y, F_Z(z)=\int_{x+y\le z}f(x,y)\,\mathrm{d}x\,\mathrm{d}y=$

$$\begin{cases}0 & \text{if } z\le -1\\ \frac{(1+a)^2}{2} & \text{if } -1<z<0\\ 1-\frac{(1-a)^2}{2} & \text{if } 0\le z<1\\ 1 & \text{if } z\ge 1\end{cases}
\quad
\begin{cases}f_Z(z)= & \\ 1+z, & -1<z<0\\ 1-z, & 0\le z<1\\ 0, & \text{otherwise}\end{cases}$$

3.期望

$\mu_X=E(X)=\int x\,\mathrm{d}F(x)<\infty$, 称期望存在。
懒惰统计学家法则 $E(r(X))=\int r(x)dF_X(x)$.
k 阶矩 $E(X^k)$, 中心矩 $E((X-\mu)^k)$。
$\sigma^2=V(X)=E((X-\mu)^2)=E(X^2)-E(X)^2$
样本 $\bar{X}=1/n\sum X_i, S_n^2=1/(n-1)\sum(X_i-\bar{X})^2$
定理：$E(\bar{X}_n)=\mu, V(\bar{X}_n)=\sigma^2/n, E(S_n^2)=\sigma^2$.
$Cov(X,Y)=E((X-\mu_X)(Y-\mu_Y))=E(XY)-E(X)E(Y)$，不相关：$Cov(X_i,X_j)=0$，此时两者方差可加，但不代表独立（例：$X=U(-1,1), Y=|X|$）
矩母函数 $\phi_X(t)=E(e^{tX})$。如果 $Y=aX+b$，则 $\phi_Y(t)=e^{bt}\phi_X(at)$。

test

test