概率论 12/13/2018

事件: $A \subseteq F$, $A \subset \Omega$

独立: A与B独立 $\Leftrightarrow P(A\cap B)=P(A)P(B)$

$A_1 \dots A_n$ 独立: $\forall 1\le i_1<\dots<i_k\le n$, $P(A_{i_1}\cap\dots\cap A_{i_k})=P(A_{i_1})\cdots P(A_{i_k})$

定义:
- 样本空间: Ω 集合
- 概率空间:
  - σ代数 F: 以 Ω 的子集构成的集合 (对补, 可数并运算封闭)
    - 有限、可数: 由所有 Ω 子集构成
    - 不可数: 2^Ω 太复杂
  - 测度 P:
    - 非负: $\forall A\in F$, $P(A)\ge 0$
    - 零测: $P(\emptyset)=0$
    - 可数可加: 任一两两互不相交的 $A_i\in F$, 有 $P(\bigcup_{n=1}^\infty A_n)=\sum_{n=1}^\infty P(A_n)$
  - 概率: $P(\Omega)=1$ 的测度 (有限可加性可推)

性质:
① $P(A\cup B)+P(A\cap B)=P(A)+P(B)$
② $A\subset B \Rightarrow P(A)\le P(B)$
③ $0\le P(A)\le 1$
④ $P(A_1\cup A_2\cup\dots\cup A_n)\le\sum P(A_k)$
取等: $A_1\dots A_n$ 均互不相交

事件 → 独立; 条件概率与 Bayes

条件概率: $P(A|S)=\frac{P(AS)}{P(S)}$

性质:
① 连乘: $P(A_1\dots A_n)=P(A_1)P(A_2|A_1)\cdots P(A_n|A_{n-1}\dots A_1)$
② 全概率: 若 $\Omega=\sum_n A_n$ 为 Ω 的划分 (互不相交), 则 $\forall B$, $P(B)=\sum_n P(A_n)P(B|A_n)$

Bayes: $P(A_n|B)=\frac{P(A_n)P(B|A_n)}{\sum_k P(A_k)P(B|A_k)}$

应用: 由果推因

随机变量与分布

随机变量: $X(\omega)$, $\omega\in\Omega$, $\forall x$, $\{\omega: X(\omega)\le x\}\in F$ (F代数)
(Ω 为 X 定义域, 值域为实数)

$\varphi(X,Y)$ 也是随机变量 ($+,-,\times,\div\dots$) 性质

概率分布函数 (离散): $P(X=k)$, $\sum_k P(X=k)=1$

密度函数 (连续): $f(u)\ge 0$, $\int_{-\infty}^{\infty} f(u)du=1$
(联合: $f(u,v)$)

离散: $P(X=k)=F_X(k)-F_X(k-1)$
连续: $f(u_0)=F_X'(u_0)$, $F_X(u_0)=\int_{-\infty}^{u_0}f(x)dx$

分布函数 (离散连续): $P(X(\omega)\le x)=F_X(x)$
性质: $F(-\infty)=0$, $F(\infty)=1$, 右连续

随机变量的独立: 对 $X_1\dots X_n$ 独立:
- 一般: 事件 $\{X_1\le x_1\}\dots\{X_n\le x_n\}$ 独立, 对 $\forall x_1\dots x_n$ 有 $F(x_1\dots x_n)=F_1(x_1)\cdots F_n(x_n)$
- 可数值: $P(X_1=x_1,\dots,X_n=x_n)=P(X_1=x_1)\cdots P(X_n=x_n)$
- 不可数值: $f(u_1,\dots,u_n)=f_1(u_1)\cdots f_n(u_n)$

可数值性质:
① $\forall S_1\dots S_n$, 独立: $P(X_1\in S_1,\dots,X_n\in S_n)=P(X_1\in S_1)\cdots P(X_n\in S_n)$
② $\forall$ 实值函数 $\varphi_1(X_1),\dots,\varphi_n(X_n)$ 独立

生成函数, 特征函数, 矩母函数
(用生成函数研究分布) (一一对应)

矩母函数: Laplace 变换: (任意) $E(e^{\lambda X})=\int_{-\infty}^{\infty}e^{\lambda u}f(u)du$ $\left(\sum_k e^{tk}P(X=k)\right)$

k阶矩: $\frac{d^k}{dt^k}M(t)\big|_{t=0}=E(X^k)$

特征函数: Fourier 变换 $E(e^{i\theta X})=\int_{-\infty}^{\infty}e^{i\theta u}f(u)du$ $\left(\sum_k e^{itk}P(X=k)\right)$

性质:
① 与概率分布一一对应 (同矩母/特征 ⇒ 同分布)
② 若 $X_1\dots X_n$ 独立, 则 $X=X_1+\dots+X_n$ 的矩母/特征:
$E(e^{\lambda X})=E(e^{\lambda X_1})\cdots E(e^{\lambda X_n})$
$E(e^{i\theta X})=E(e^{i\theta X_1})\cdots E(e^{i\theta X_n})$

总结: 一一对应 (矩母/生成/特征与分布); 加/乘变换

生成函数 (离散)

$g(z)=\sum_{j=0}^\infty P(X=j)z^j$

$|z|\le 1$: 收敛, $g(1)=1$

应用: $E(X)=g'(1)$, $E(X^2)=g'(1)+g''(1)$

性质:
① $\sigma^2(X)=g'(1)+g''(1)-(g'(1))^2$
② 与概率分布一一对应 (同生成函数 ⇒ 同分布)
③ 若 $X_1\dots X_n$ 独立, 则 $X_1+\dots+X_n$ 生成函数 $g=g_1\cdots g_n$

卷积: $c_i=\sum_{j=0}^i a_j b_{i-j}$
$=P(X+Y=i)$
$=g_1g_2$ 中 z^i 系数

期望, 方差, 相关

期望:
- 离散: $E(X)=\sum_{\omega\in\Omega}X(\omega)P(\{\omega\})$
- 连续: $E(X)=\int_{-\infty}^{\infty}u f(u)du$
- $E(\varphi(X,Y))=\iint\varphi(x,y)f(x,y)dxdy$

一般: $E(\varphi(X))=\sum_{\omega\in\Omega}\varphi(X)P(\{\omega\})$

若 $\varphi(X)=X^r$: r阶矩; r阶中心矩: $\varphi(X)=(X-E(X))^r$

性质:
① $\forall X_1\dots X_n$, $E(X_1+\dots+X_n)=E(X_1)+\dots+E(X_n)$
② $E(I_A)=P(A)$
③ $P(A_1\cup\dots\cup A_n)=\sum P(A_i)-\sum P(A_iA_j)+\dots$ (容斥原理)

方差与偏差:
$\sigma^2(X)=E((X-E(X))^2)$
$=E(X^2)-[E(X)]^2$
$\sigma(X)=\sqrt{\sigma^2(X)}$

性质: 若 $X_1\dots X_n$ 独立, σ^2 有限, 则 $\sigma^2(X_1+\dots+X_n)=\sigma^2(X_1)+\dots+\sigma^2(X_n)$

协方差: $E((X-EX)(Y-EY))=Cov(X,Y)=E(XY)-E(X)E(Y)$

相关系数: $\rho(X,Y)=\frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$

性质定义:
① $-1\le\rho\le 1$ (Cauchy-Schwarz 不等式) $[E(XY)]^2\le E(X^2)E(Y^2)$
② 若 X, Y 独立, $Cov(X,Y)=\rho(X,Y)=0$ (不相关)
④ 若 X, Y 独立, $E(XY)=E(X)E(Y)$

集合论

$A^c=\Omega\setminus A$, $A\cup B$ (并; $A+B$), $A\setminus B$ (差) ($B\subset A$: $A-B$)
$A\cap B$
$A\triangle B$ 对称差
$\bigcup_{n=1}^\infty$, $\bigcap_{n=1}^\infty$

指示函数 $I_A(\omega)=\begin{cases}1 & \omega\in A\\ 0 & \omega\notin A\end{cases}$

$a\vee b=\max(a,b)$
$a\wedge b=\min(a,b)$

$I_{A\cap B}=I_A\times I_B=I_A\wedge I_B$
$I_{A\cup B}=I_A\vee I_B$
$I_{A\cap B}+I_{A\cup B}=I_A+I_B$

分布 12/25/2018

名称	密度/概率/母函数/分布函数	μ, σ^2, 矩母	矩母/特征函数	应用场合与性质
伯努利分布	$P(X=1)=p$, $P(X=0)=1-p$	$\mu=p$, $\sigma^2=p(1-p)=pq$	$f(z)=q+pz$	e.g. 掷一枚硬币
二项分布 $X\sim B(n,p)$	$P(S_n=k)=f(k;n,p)=\binom{n}{k}p^k(1-p)^{n-k}$	$E(S_n)=np$, $\sigma^2(S_n)=npq$	$f(z)=(q+pz)^n$	e.g. 掷n枚硬币，正面次数；若 $X\sim B(n,p)$，$Y\sim B(m,p)$，X,Y独立，则 $X+Y\sim B(n+m,p)$；极限情况：取 $p=\frac{\alpha}{n}$，$\lim_{n\to\infty} A_k(n,k)=\pi_k(\alpha)=\frac{\alpha^k}{k!}e^{-\alpha}$ → Poisson分布
Poisson分布 $\pi(\alpha)$	$P(X=k)=\frac{e^{-\alpha}}{k!}\alpha^k \ (k\in N)$；n较大时，Poisson分布可用 $\alpha=np$ 近似二项分布	$E(X)=\alpha$, $\sigma^2(X)=\alpha$	$f(z)=e^{\alpha(z-1)}$, $M(t)=e^{\alpha(e^t-1)}$, $\varphi(w)=e^{\alpha(e^{iw}-1)}$	① 若 $X_1\ldots X_j$ 独立，均满足 $\pi(\alpha_1)\ldots\pi(\alpha_j)$，则 $X_1+\ldots+X_j\sim\pi(\alpha_1+\ldots+\alpha_j)$；② stirling公式 $\lim_{n\to\infty}\frac{n!e^n}{n^{n+1/2}}=\sqrt{2\pi}$；③ 李店：泊松分布可逼近二项分布 ($p_n$ 与n有关, $p_n=\frac{\alpha}{n}$)，$\alpha=np$

正态分布 $N(\mu,\sigma^2)$

密度：$\varphi(x)=\frac{1}{\sqrt{2\pi}}\ldots$

一般：$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$F(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{x}e^{-\frac{(u-\mu)^2}{2\sigma^2}}du$

标准 $\left(\frac{x-\mu}{\sigma}\right)$：$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

$\mu=\mu$, $\sigma^2=\sigma^2$；标准：$\mu=0,\sigma=1$

$M(\theta)=\int_{-\infty}^{\infty}e^{\theta u}\varphi(u)du=e^{\frac{\theta^2}{2}}$（矩母函数）

$M(\theta)$（一般）$=e^{\mu\theta+\frac{\sigma^2\theta^2}{2}}$

性质：① 对称 $\varphi(x)=\varphi(-x)$（标准）

② 若 X_j 独立，$X_j\sim N(m_j,\sigma_j^2)$，则 $X_1+\ldots+X_n$ 服从 $N(m_1+\ldots+m_n,\sigma_1^2+\ldots+\sigma_n^2)$

③ 中心极限定理：若 $S_n=X_1+\ldots+X_n$ 独立同分布，$E(X)=m$，$Var(X)=\sigma^2$，则有 $\lim_{n\to\infty}P\left(a<\frac{S_n-E(S_n)}{\sigma(S_n)}\le b\right)=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx$（大量独立同分布 → 正态分布）$\left(a<\frac{S_n-n\mu}{\sqrt{n}\sigma}\le b\right)$

④ 大数定律：若 S_n 满足中心极限定理，则 $\forall c$，$\lim_{n\to\infty}P\left(\left|\frac{S_n}{n}-m\right|<c\right)=1$（正态分布：限制了取值，同因）

⑤ Chebyshev不等式：若X二阶矩有限，则 $P(|X|\ge c)\le\frac{E(X^2)}{c^2}$ $(\forall c>0)$

卡方分布

统计应用：若总体正态分布 $N(\mu,\sigma^2)$，则 $Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ 服从标准正态分布

$\chi^2=X_1^2+\ldots+X_n^2$，其中 $X_1\ldots X_n$ 独立，服从标准正态分布，

则 $f(x)=\frac{1}{2^{\frac{v}{2}}\Gamma(\frac{v}{2})}x^{\frac{v}{2}-1}e^{-\frac{x}{2}}$ $(x>0)$

$\Gamma(x)=\int_0^{\infty}t^{x-1}e^{-t}dt$

$\mu=v$（自由度，变量个数），$\sigma^2=2v$

$M(t)=(1-2t)^{-\frac{v}{2}}$；$\varphi(w)=(1-2iw)^{-\frac{v}{2}}$

性质：① $X_1^2+\ldots+X_n^2$ 服从自由度 $v=n$ 卡方分布（标准正态）

② $W=U_1+\ldots+U_k$：$V_1+\ldots+V_k$ 卡方分布（添加也独立）（$U_1\ldots U_k$ 独立，V_i 卡方分布）

③ 统计应用：若总体正态分布，则 $Z=\frac{nS^2}{\sigma^2}$：$v=n-1$ 卡方分布

t分布

$f(t)=\frac{\Gamma(\frac{v+1}{2})}{\sqrt{v\pi}\,\Gamma(v/2)}\left(1+\frac{t^2}{v}\right)^{-\frac{v+1}{2}}$

$\mu=0$（对称），$\sigma^2=\frac{v}{v-2}$ $(v>2)$

$\Gamma(x)=\int_0^{\infty}t^{x-1}e^{-t}dt$

性质：① 正态、卡方、t：若Y为标准正态，Z为v卡方分布，则 $T=\frac{Y}{\sqrt{Z/v}}$：自由度v的T分布

② 统计应用：若总体正态分布，则 $T=\frac{\bar{X}-\mu}{\sqrt{S^2/(n-1)}}$：自由度n-1的T分布

③ 特殊：$v=1$：Cauchy分布；$v\to\infty$：标准正态分布

Cauchy分布

$f(t)=\frac{a}{\pi(x^2+a^2)}$，有能但无 μ,σ^2

Gamma分布 $(0,\infty)$：$f(x)=\frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha\Gamma(\alpha)}$，$\mu=\alpha\beta$，$\sigma^2=\alpha\beta^2$

Beta分布 $(0,1)$：$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$ $\left(B(\alpha,\beta)=\int_0^1x^{\alpha-1}(1-x)^{\beta-1}=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}\right)$

$\mu=\frac{\alpha}{\alpha+\beta}$，$\sigma^2=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

超几何分布

设b黑球，r红球，X_n：n抽球后黑球个数（不放回）

$P(X_n=k)=\frac{\binom{b}{k}\binom{r}{n-k}}{\binom{b+r}{n}}$

$\mu=np$，$\sigma^2=npq\frac{N-n}{N-1}$

$N\to\infty$：趋于二项分布

几何分布

Bernoulli试验：第一次成功（第k次）

$P(X=k)=pq^{k-1}$

$\mu=\frac{1}{p}$，$\sigma^2=\frac{q}{p^2}$

$M(t)=\frac{pe^t}{1-qe^t}$

Pascal分布

Bernoulli试验：第r次成功

$P(X=k)=\binom{k-1}{r-1}p^rq^{k-r}$

$\mu=\frac{r}{p}$，$\sigma^2=\frac{rq}{p^2}$

$M(t)=\left(\frac{pe^t}{1-qe^t}\right)^r$（相当于r个独立几何分布之和）

多项分布

$A_1\ldots A_k$ 的k项分布 $(p_1+\ldots+p_k=1)$

$P(X_1=n_1,\ldots,X_k=n_k)=\frac{n!}{n_1!\ldots n_k!}p_1^{n_1}p_2^{n_2}\ldots p_k^{n_k}$

指数分布

$f(x)=\begin{cases}\alpha e^{-\alpha x} & x>0\\ 0 & x\le 0\end{cases}$

$\mu=\frac{1}{\alpha}$，$\sigma^2=\frac{1}{\alpha^2}$，$M(t)=\frac{\alpha}{\alpha-t}$

性质：① $\varphi(s+t)=\varphi(s)\varphi(t)$

② Markov性：$P(T>s+t\mid T>s)=P(T>t)$

Statistics
12/27/2018

↓

总体与样本

总体：总体属性：μ, σ^2 等（相对论）

样本：抽取的总体一部分 | 随机变量：有 $\bar x$, s 才能抽取，随抽取变
| 确定性（$\bar x$, s 等可求）

↓

统计量：均值 $\bar x = \frac{x_1+x_2+\dots+x_n}{n}$

方差 $s^2 = \frac{(x_1-\bar x)^2+\dots+(x_n-\bar x)^2}{n}$ （$S^2 = \frac{ns^2}{n-1}$）

① $E(\bar X) = \mu_{\bar x} = \mu$（$\bar X$为$\mu$的无偏估计）

② $\sigma_{\bar x}^2 = E((\bar X-\mu)^2) = \frac{\sigma^2}{n}$（放回）

$\frac{\sigma^2}{n}\frac{N-n}{N-1}$（不放回）

（e.g. 比例分布：$E(p)=p$, $\sigma(p)=\sqrt{\frac{pq}{n}}$）

③ 若总体服从 $N(\mu,\sigma^2)$，

则 $\frac{\bar X-\mu}{\sigma/\sqrt n}$ 服从标准正态分布

④ 差与和两个总体的样本 S_1, S_2：

$\mu_{S_1\pm S_2} = \mu_{S_1}\pm\mu_{S_2}$, $\sigma_{S_1\pm S_2} = \sqrt{\sigma_{S_1}^2+\sigma_{S_2}^2}$

① $\mu_{s^2} = \frac{n-1}{n}\sigma^2$

$\mu_{S^2}=E(S^2)$

$\sigma^2 = \mu_{S^2} = E(S^2)$（$S^2$为$\sigma^2$无偏估计）

② 若总体服从 $N(\mu,\sigma^2)$

则 $\frac{ns^2}{\sigma^2}$, $\frac{(n-1)S^2}{\sigma^2}$ 服从自由度 $\nu=n-1$ 卡方分布

③ 总体服从 $N(\mu,\sigma^2)$

则 $\frac{\bar X-u}{S/\sqrt n}$, $\frac{\bar X-u}{s/\sqrt{n-1}}$ 服从自由度 $\nu=n-1$ t分布

↓

估计：样本→总体

无偏估计：$E(估计)=\mu$

有效估计：同样均值，方差越小越有效

点估计：估值　区间估计：估区间

（集合）置信区间：

让总体参数在区间内概率 $=P_0$ 的区间

做法：①：找抽样分布：利用样本总体参数与已知分布（$\bar x$, μ 已知… ; s, σ 已知… ; $\bar x$, s, u → t）

② 若为正态/t分布：求 $F^{-1}(\frac{1+P}{2})$ → 临界值 z

卡方分布：求 $F^{-1}(\frac{1-P}{2})$, $F^{-1}(\frac{1+P}{2})$ → 临界值 2

二项分布（比例）→ 正态近似

③：代回，求区间端点（二项分布：置信区间为 $P\pm z_c\sigma_P$）

$\sigma_P=\sqrt{\frac{pq}{n}}$

（和与差：$S_1\pm S_2 \pm z_c\sqrt{\sigma_{S_1}^2+\sigma_{S_2}^2}$

（方差比 $F_{\frac{1-P}{2}} < \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < F_{\frac{1+P}{2}}$

自由度为 $m-1$, $n-1$
S_1　S_2

↓

极大似然：

Fisher方法：

① 联合密度函数

$L = f(x_1,\theta)\cdots f(x_n,\theta)$

（$x_1\cdots x_n$ 为样本

θ：待估量　f：总体密度函数）

② 取对数求导：

$\frac{1}{L}\frac{\partial L}{\partial\theta} = \sum \frac{\frac{\partial f(x_i,\theta)}{\partial\theta}}{f(x_i,\theta)}$

③ 取 $\frac{\partial L}{\partial\theta}=0$，解 $\sum\frac{\frac{\partial f(x_i,\theta)}{\partial\theta}}{f(x_i,\theta)} = 0$

↓

解出 $\theta(x_1\cdots x_n)$ 为极大似然估计

⇓

e.g. 总体正态分布：$\mu=\bar X$, $\sigma^2 = s^2$

Bayes估计（后验估计）

（估计）

实际：先验分布 $\pi(\theta)$（事先均匀分布）

联合密度：$h(x,\theta)=\pi(\theta)f(x|\theta)$

求法：①：求联合密度函数 $h(x,\theta)=\pi(\theta)f(x|\theta)$

若X二项分布：$h(x,\theta)=\pi(\theta)\theta^x(1-\theta)^{n-x}\binom{n}{x}$

（若多次 $h(x,\theta)=\pi(\theta)\cdot f(x_1|\theta)\cdots f(x_n|\theta)$）

②：求后验密度

$\pi(\theta|x) = \frac{h(x,\theta)}{\int_0^1 h(x,\theta)d\theta} = k\,h(x,\theta)$

（下方与θ无关）$=\frac{1}{m(x)}$

（若X二项分布：$\int_0^1 h(x,\theta)d\theta$

$= B(x+1, n-x+1)$（β函数）

③ 求 $E(\pi(\theta|x))$ → Bayes估计值

（X：二项分布：$E=\frac{B(x+2,n-x+1)}{B(x+1,n-x+1)}$

$=\frac{x+1}{n+2}$

统计：①：求极差（最大-最小）

② 设定interval（间隔），组界

③ 统计Frequency

（样本均值分布）

不必考虑所有可能

频率分布

1. 得到数据组（…），分类：

分类：一类区间：如身高60：记为 $[59.5, 60.5)$…

区间间隔 60-62：即 $[59.5, 62.5)$，间隔3，等间隔

分类标志　区间中值（60-62：中间61）

计算：均值 $\bar x = \frac{f_1x_1+\dots+f_kx_k}{n}$ （f_i：每组频数，x_i：分类标志（中值））

方差：$s^2=\frac{f_1(x_1-\bar x)^2+\dots+f_k(x_k-\bar x)^2}{n}$

r阶矩：均值 $m_r = \frac{f_1(x_1-\bar x)^r+\dots+f_k(x_k-\bar x)^r}{n}$

$m_r' = \frac{f_1x_1^r+\dots+f_kx_k^r}{n}$