概率论与数理统计小班辅导

杨维铠

2017 年 5 月 28 日

1. 概率初步

Bayes 公式

$$P(XY) = P(X)P(Y|X) = P(Y)P(X|Y)$$

Bayes 公式

$$P(XY) = P(X)P(Y|X) = P(Y)P(X|Y)$$

例子

小明在紫操踢球时将球踢出操场，已知球落在草丛的几率是 $\frac{2}{3}$，落在马路上的几率是 $\frac{1}{3}$。而如果球落在了马路上，小明有 $\frac{9}{10}$ 的几率一眼找到。现在小明往马路上瞄了一眼没有找到，问球落在马路上的概率是多少。

Bayes 公式

证明.

记 X 为球在马路上的概率，Y 为小明没在马路上找到球的概率。则
$P(X) = \frac{1}{3}, P(Y|X) = \frac{1}{10}, P(Y) = \frac{2}{3} + \frac{1}{3} \times \frac{1}{10}$，立得
$P(X|Y) = \frac{1}{21}$ $\square$

2. 随机变量

常见分布

名称	形式	期望	方差
二项分布	$C_n^i p^i (1-p)^{n-i}$	np	$np(1-p)$
泊松分布	$e^{-\lambda}\frac{\lambda^i}{i!}$	λ	λ
正态分布	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
均匀分布	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布	$\lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

例子

蝉的产卵数服从 $P(\lambda)$，卵发育成成虫概率为 p 且独立。求后代数的分布。

例子

蝉的产卵数服从 $P(\lambda)$，卵发育成成虫概率为 p 且独立。求后代数的分布。

证明.

记后代数为 X，则

$$
\begin{aligned}
P(X=i) &= \sum_{j=0}^{\infty} e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}\frac{(i+j)!}{i!\ j!}p^i(1-p)^j \\
&= e^{-\lambda}\frac{(\lambda p)^i}{i!}\sum_{j=0}^{\infty}\frac{(\lambda(1-p))^j}{j!} \\
&= e^{-\lambda}\frac{(\lambda p)^i}{i!}\times e^{\lambda(1-p)} \\
&= e^{-\lambda p}\frac{(\lambda p)^i}{i!}
\end{aligned}
$$

□

边际分布和条件分布

借助离散型理解连续型

边际分布和条件分布

借助离散型理解连续型

边际分布 $f_X(x) = \int_{-\infty}^{\infty} f(x, y)\mathrm{d}y$

边际分布和条件分布

借助离散型理解连续型

边际分布 $f_X(x) = \int_{-\infty}^{\infty} f(x, y)\mathrm{d}y$

条件分布 $f_{X|Y}(x|y) = f_X(x|Y = y) = \frac{f(x,y)}{f_Y(y)}$

求密度函数

可逆映射：密度变换法

求密度函数

可逆映射：密度变换法

$$\iint_B f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 =$$

$$\iint_A f(h_1(y_1, y_2), h_2(y_1, y_2))|J|\mathrm{d}y_1\mathrm{d}y_2$$

其中 $J = \begin{bmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{bmatrix}$.

求密度函数

可逆映射：密度变换法

$$\iint_B f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 =$$

$$\iint_A f(h_1(y_1, y_2), h_2(y_1, y_2))|J|\mathrm{d}y_1\mathrm{d}y_2$$

其中 $J = \begin{bmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{bmatrix}.$

不可逆映射：直接法

求密度函数

可逆映射：密度变换法

$$\iint_B f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 =$$

$$\iint_A f(h_1(y_1, y_2), h_2(y_1, y_2))|J|\mathrm{d}y_1\mathrm{d}y_2$$

其中 $J = \begin{bmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{bmatrix}$.

不可逆映射：直接法

$$P(Y \leq y) = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)\mathrm{d}x = \int_{-\infty}^{y} l(y)\mathrm{d}y$$

$$l(y) = \frac{1}{2\sqrt{y}}[f(\sqrt{y}) + f(-\sqrt{y})]$$

求密度函数

例子

给定随机变量 Y，其 cdf 为 $F(y)$，可逆，试通过均匀分布 $X \sim U(0,1)$，构造出映射 g，使 $Y = g(X)$

求密度函数

例子

给定随机变量 Y，其 cdf 为 $F(y)$，可逆，试通过均匀分布 $X \sim U(0,1)$，构造出映射 g，使 $Y = g(X)$

证明.

$F(y) = P(Y < y) = P(g(X) < y) = P(X < g^{-1}(y)) = g^{-1}(y)$ □

数字特征

数字特征

- 期望：线性性质，若独立则有
  $E(XY) = E(X)(Y)$

数字特征

- 期望：线性性质，若独立则有
  $E(XY) = E(X)(Y)$
- 方差：常用 $E(X^2) - E(X)^2$，若独立则有
  $Var(X + Y) = Var(X) + Var(Y)$

数字特征

- 期望：线性性质，若独立则有
  $E(XY) = E(X)(Y)$
- 方差：常用 $E(X^2) - E(X)^2$，若独立则有
  $Var(X + Y) = Var(X) + Var(Y)$
- 条件期望：条件分布下的期望，满足重期望公式

数字特征

- 期望：线性性质，若独立则有
  $E(XY) = E(X)(Y)$
- 方差：常用 $E(X^2) - E(X)^2$，若独立则有
  $Var(X + Y) = Var(X) + Var(Y)$
- 条件期望：条件分布下的期望，满足重期望公式
- 协方差：$cov(X, Y) = E(XY) - E(X)E(Y)$

数字特征

- 期望：线性性质，若独立则有

  $E(XY) = E(X)(Y)$

- 方差：常用 $E(X^2) - E(X)^2$，若独立则有

  $Var(X + Y) = Var(X) + Var(Y)$

- 条件期望：条件分布下的期望，满足重期望公式
- 协方差：$cov(X, Y) = E(XY) - E(X)E(Y)$
- 线性相关系数：$\frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

数字特征

例子

$f(x,y)=2(0<y<x<1)$，分析 X,Y 是否独立，并求 $P(0.25\le Y\le 0.5|X=0.75)$，并求 $X=a$ 时 Y 在均方误差意义下的最优估计。

数字特征

例子

$f(x,y)=2(0<y<x<1)$，分析 X,Y 是否独立，并求 $P(0.25\le Y\le 0.5|X=0.75)$，并求 $X=a$ 时 Y 在均方误差意义下的最优估计。

证明.

$$(1)E(X)=\frac{2}{3},E(Y)=\frac{1}{3},E(XY)=\frac{1}{4}$$
$$(2)f_X(x)=2x,f_Y(y)=2(1-y),f(y|X=a)=\frac{1}{a}$$
$$(3)E((Y-c)^2)=E(Y^2)-2cE(Y)+c^2,c=E(Y)=\frac{a}{2}\quad\square$$

大数定律

大数定律

- $P(Y \geq \epsilon) \leq \frac{E(Y)}{\epsilon}$

大数定律

- $P(Y \geq \epsilon) \leq \frac{E(Y)}{\epsilon}$
- $P(Y - E(Y) \geq \epsilon) \leq \frac{Var(Y)}{\epsilon^2}$

大数定律

- $P(Y \geq \epsilon) \leq \frac{E(Y)}{\epsilon}$
- $P(Y - E(Y) \geq \epsilon) \leq \frac{Var(Y)}{\epsilon^2}$
- $P(\frac{X-\mu}{\sigma/\sqrt{n}} \leq x) = \Phi(x)$

3. 参数估计

点估计

点估计

- 矩估计

点估计

- 矩估计
- 极大似然估计

点估计

- 矩估计
- 极大似然估计
- 中位数估计

点估计

- 矩估计
- 极大似然估计
- 中位数估计
- 无偏性 $E_{\theta_1,...,\theta_k}[\hat{g}(X_1,...,X_n)] = g(\theta_1,...,\theta_k)$

点估计

- 矩估计
- 极大似然估计
- 中位数估计
- 无偏性 $E_{\theta_1,\ldots,\theta_k}[\hat{g}(X_1,\ldots,X_n)] = g(\theta_1,\ldots,\theta_k)$
- 相合性

$$\lim_{n\to\infty} P_{\theta_1,\ldots,\theta_k}(|T(X_1,\ldots,X_n) - g(\theta_1,\ldots,\theta_k)| \geq \epsilon) = 0$$

例子

$f(x) = \frac{\theta}{x^{\theta+1}}(x > 1)$，求矩估计、极大似然估计，并证明 $\frac{1}{\theta^*}$ 是 $\frac{1}{\theta}$ 的无偏估计。

例子

$f(x)=\frac{\theta}{x^{\theta+1}}(x>1)$，求矩估计、极大似然估计，并证明 $\frac{1}{\theta^*}$ 是 $\frac{1}{\theta}$ 的无偏估计。

证明.

$(1)\frac{\theta}{\theta-1}=\bar{X}$

$(2)\theta^*=\frac{n}{ln(X_1X_2...X_n)}$

(3) 注意到 $E(ln(X))=\frac{1}{\theta}$ $\square$

三大分布

三大分布

- χ_n^2 分布：n 个独立标准正态分布的和

三大分布

- χ_n^2 分布：n 个独立标准正态分布的和
- t_n 分布：$\dfrac{N}{\sqrt{\chi_n^2/n}}$

三大分布

- χ_n^2 分布：n 个独立标准正态分布的和
- t_n 分布：$\dfrac{N}{\sqrt{\chi_n^2/n}}$
- $F_{m,n}$ 分布:$\dfrac{\chi_m^2/m}{\chi_n^2/n}$

三大分布

- χ_n^2 分布：n 个独立标准正态分布的和
- t_n 分布：$\frac{N}{\sqrt{\chi_n^2/n}}$
- $F_{m,n}$ 分布:$\frac{\chi_m^2/m}{\chi_n^2/n}$
- $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

三大分布

- χ_n^2 分布：n 个独立标准正态分布的和
- t_n 分布：$\frac{N}{\sqrt{\chi_n^2/n}}$
- $F_{m,n}$ 分布:$\frac{\chi_m^2/m}{\chi_n^2/n}$
- $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- $\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$

三大分布

三大分布

- σ^2 已知，$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

三大分布

- σ^2 已知，$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- σ^2 未知，$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$，$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$

三大分布

- σ^2 已知，$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- σ^2 未知，$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$，$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$
- $\sigma_1^2 = \sigma_2^2$，则 $\bar{X} \sim N(\mu_1, \sigma^2/n), \bar{Y} \sim N(\mu_2, \sigma^2/m)$ 故 $\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \sigma^2(\frac{1}{n} + \frac{1}{m}))$

三大分布

- σ^2 已知，$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- σ^2 未知，$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$，$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$
- $\sigma_1^2 = \sigma_2^2$，则 $\bar{X} \sim N(\mu_1, \sigma^2/n), \bar{Y} \sim N(\mu_2, \sigma^2/m)$ 故 $\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \sigma^2(\frac{1}{n} + \frac{1}{m}))$
- 又 $\frac{(n-1)S_1^2}{\sigma^2} \sim \chi^2_{n-1}, \frac{(m-1)S_2^2}{\sigma^2} \sim \chi^2_{m-1}$，故 $\frac{(n-1)S_1^2+(m-1)S_2^2}{\sigma^2} \sim \chi^2_{n+m-2}$，仿上述操作亦可得 t 分布

区间估计

区间估计

- $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，得 $\mu \in \bar{X} \pm u_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

区间估计

- $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，得 $\mu \in \bar{X} \pm u_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
- $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ 得 $\mu \in \bar{X} \pm t_{n-1}(\frac{\alpha}{2}) \frac{S}{\sqrt{n}}$

区间估计

- $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，得 $\mu \in \bar{X} \pm u_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
- $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ 得 $\mu \in \bar{X} \pm t_{n-1}(\frac{\alpha}{2})\frac{S}{\sqrt{n}}$
- $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ 得 $\sigma^2 \in [\frac{(n-1)S^2}{\chi^2_{n-1}(\alpha/2)}, \frac{(n-1)S^2}{\chi^2_{n-1}(1-\alpha/2)}]$

区间估计

- $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$，得 $\mu \in \bar{X} \pm u_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
- $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ 得 $\mu \in \bar{X} \pm t_{n-1}(\frac{\alpha}{2}) \frac{S}{\sqrt{n}}$
- $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ 得 $\sigma^2 \in [\frac{(n-1)S^2}{\chi^2_{n-1}(\alpha/2)}, \frac{(n-1)S^2}{\chi^2_{n-1}(1-\alpha/2)}]$
- 大样本：S 直接代替 σ^2

选举问题

$$\frac{\bar{p}-p}{\sqrt{p(1-p)/n}} \overset{\text{近似}}{\sim} N(0,1)$$

选举问题

$$\frac{\bar{p}-p}{\sqrt{p(1-p)/n}} \overset{\text{近似}}{\sim} N(0,1)$$

可以用 $S, m, \bar{p}(1-\bar{p}), 0.25$ 来近似 $p(1-p)$

选举问题

$$\frac{\bar{p}-p}{\sqrt{p(1-p)/n}} \overset{近似}{\sim} N(0,1)$$

可以用 $S, m, \bar{p}(1-\bar{p}), 0.25$ 来近似 $p(1-p)$

例子

需要多少样本可以有 99% 的把握得到误差不超过 3% 的预测结果？

4. 假设检验

临界值检验 &P 值检验

临界值检验 &P 值检验

- H_0, H_1 错误的认识

临界值检验 &P 值检验

- H_0, H_1 错误的认识
- 临界值检验：与区间估计对偶

临界值检验 &P 值检验

- H_0, H_1 错误的认识
- 临界值检验：与区间估计对偶
- P 值检验：比当前情况更极端的概率

拟合优度检验

$$\sum_{i=0}^{k} \frac{(O_i - e_i)^2}{e_i} \sim \chi^2_{k-1}$$

例子

唐老师让小明扔 600 次骰子，得到结果是 98,101,102,103,96,100。问小明是否有伪造数据的嫌疑。

拟合优度检验

$\sum_{i=0}^{k} \frac{(O_i - e_i)^2}{e_i} \sim \chi^2_{k-1}$

例子

唐老师让小明扔 600 次骰子，得到结果是 98,101,102,103,96,100。问小明是否有伪造数据的嫌疑。

证明.

计算得其卡方值为 0.34，而 $\chi^2_5(0.995) = 0.412$，也就是说小明的结果好到 1000 次实验里平均成功次数不到 5 次，因此有理由怀疑其造假。