习题答案

司继春

上海对外经贸大学统计与信息学院

1 概率

练习 1. 整数集 $\mathbb{Z}$ 是可数集还是不可数集? 有理数集 $\mathbb{Q}$ 是可数集还是不可数集? 开区间 $(0,1)$ 是可数集还是不可数集?

解答. 整数集 $\mathbb{Z}$ 是可数集，有理数集 $\mathbb{Q}$ 是可数集，开区间 $(0,1)$ 是不可数集。

练习 2. 思考题: 标准篮球的直径为 $24.6cm$，而标准篮筐的直径为 $45cm$，如果篮球垂直入框，中心落点均匀的落在篮筐内，请问投出空心球的概率有多大? 如果篮球以 $60°$ 角入框呢? 篮球以多大的角度入框则不可能投出空心球?

练习 3. 试用交、并、补三个运算表示 $(E\triangle F)^c$。

解答. $E\Delta F = (E\backslash F)\cup(F\backslash E) = (E\cap F^c)\cup(F\cap E^c)$

练习 4.

1. 上例中，包含 $\{\clubsuit\}$ 的最小 σ-代数是?
2. 现在抛一枚骰子，则结果的样本空间为 $\Omega_4=\{1,2,3,4,5,6\}$，那么包含所有单个样本点的 σ-代数 $\mathscr{F}_4$ 中有多少个事件?

解答. 1. $\{\emptyset,\Omega,\{\clubsuit\},\{\heartsuit,\spadesuit,\diamondsuit\}\}$ 2. $\begin{pmatrix}6\\0\end{pmatrix}+\begin{pmatrix}6\\1\end{pmatrix}+\begin{pmatrix}6\\2\end{pmatrix}+\cdots+\begin{pmatrix}6\\6\end{pmatrix}=2^6=64$

练习 5. 百年一遇的自然灾害（即每年发生的概率为 $p=1\%$）10 年期间至少发生 1 次的概率是多少?（假设这种自然灾害每年发生与否是独立的）

解答. $1-(1-1\%)^{10}\approx 9.6\%$

练习 6. 七个人玩桌游「炸碉堡」，经过第一轮投票，已知第一轮三个行动人中有一个是坏人，剩下的四个人中也有一个是坏人。如果在第二轮中由你选择三个人作为行动人，你的目标是尽可能的选出三个好人。你有如下三个策略:

1. 从三个人中选一个，另外四个人中选两个

2. 从三个人中选两个，另外四个人中选一个
3. 完全从四个人中选出新的三个

请问以上三个策略中，哪一个策略选出三个好人的概率最高?

练习 7. 给定任意一个连续的分布函数 F 及由其定义的概率函数 P，$\mathbb{R}$ 上的单点集 $\{a, a\in\mathbb{R}\}$ 的概率 $P(\{a, a\in\mathbb{R}\})$ 是多少? 根据概率函数的性质，$\mathbb{R}$ 上的任意可数集的概率 $P(\{a_i, i=1,2,\dots,a_i\in\mathbb{R}\})$ 是多少? 所以概率为 0 的事件一定是不可能事件么?

解答. 由于函数 $F(x)$ 为连续函数，因而：

$$P(\{a, a\in\mathbb{R}\}) = \lim_{n\to\infty} P\left((a-\frac{1}{n}, a]\right) = \lim_{n\to\infty}\left[F(a) - F\left(a-\frac{1}{n}\right)\right] = 0$$

而由于 $\{a_i, i=1,2,\dots,a_i\in\mathbb{R}\}$ 为可数集，因而：

$$P(\{a_i, i=1,2,\dots,a_i\in\mathbb{R}\}) = \sum_{i=1}^{\infty} p(\{a_i\}) = \sum_{i=1}^{\infty} 0 = 0$$

因而概率为 0 的时间不一定就是不可能事件。

练习 8. 试证明定理 (4)。

解答. 1. $1 = \mathscr{P}(\Omega) = \mathscr{P}(A) + \mathscr{P}(A^c), \mathscr{P}(\Omega\backslash A) \geq 0 \Rightarrow \mathscr{P}(\Omega\backslash A) \leq 1$

2. $1 = \mathscr{P}(\Omega) = \mathscr{P}(A) + \mathscr{P}(A^c) \Rightarrow \mathscr{P}(A^c) = 1 - \mathscr{P}(A)$

3. $1 = \mathscr{P}(\Omega) = \mathscr{P}(\Omega) + \mathscr{P}(\emptyset) \Rightarrow \mathscr{P}(\emptyset) = 0$

4. 由于 $A\backslash B = A\cap B^c$，因而 $(A\backslash B)\cap(A\cap B) = (A\cap B^c)\cap(A\cap B) = A\cap(B^c\cap B) = \emptyset$，同时由于 $(A\backslash B)\cup(A\cap B) = (A\cap B^c)\cup(A\cap B) = A\cap(B^c\cup B) = A$，因而 $\mathscr{P}(A\backslash B) = \mathscr{P}(A) - \mathscr{P}(A\cap B)$。

由于 $(A\backslash B)\cap B = (A\cap B^c)\cap B = \emptyset$，$(A\backslash B)\cup B = (A\cap B^c)\cup B = (B\cup A)\cap(B\cup B^c) = (B\cup A)$，从而 $\mathscr{P}(B\cup A) = \mathscr{P}(A\backslash B) + \mathscr{P}(B) = \mathscr{P}(A) - \mathscr{P}(A\cap B) + \mathscr{P}(B)$。

5. 同上，$\mathscr{P}(B\backslash A) = \mathscr{P}(B) - \mathscr{P}(A\cap B)$，由于 $A\subset B$，所以 $\mathscr{P}(A\cap B) = \mathscr{P}(A)$，所以 $\mathscr{P}(A) = \mathscr{P}(B) - \mathscr{P}(B\backslash A) \leq \mathscr{P}(B)$。

6. 令 $B_1 = A_1$，对于 $i > 1$，令 $B_i = A_i\backslash\left(\bigcup_{j=1}^{i-1} A_j\right)$，从而 $B_i\cap B_j = \emptyset$，且

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

7. 由于 $C_i\cap C_j = \emptyset$，且 $\bigcup_i C_i = \Omega$，所以 $(A\cap C_i)\cap(A\cap C_j) = A\cap(C_i\cap C_j) = \emptyset$，$\bigcup_i(A\cap C_i) = A\cap(\bigcup_i C_i) = A\cap\Omega = A$，所以 $\sum_{i=1}^{\infty}\mathscr{P}(A\cap C_i) = \mathscr{P}(\bigcup_i(A\cap C_i)) = \mathscr{P}(A)$。

练习 9. 请给出两个事件至少有一个发生的概率，$\mathscr{P}(A \cup B)$ 的一个上界和一个下界。

解答. $\mathscr{P}(A \cup B) = \mathscr{P}(A) + \mathscr{P}(B) - \mathscr{P}(A \cap B)$，由于 $0 \leq \mathscr{P}(A \cap B) \leq \min(\mathscr{P}(A), \mathscr{P}(B))$，从而：

$$\mathscr{P}(A)+\mathscr{P}(B)-\min(\mathscr{P}(A), \mathscr{P}(B)) = \max(\mathscr{P}(A), \mathscr{P}(B)) \leq \mathscr{P}(A)+\mathscr{P}(B)-\mathscr{P}(A \cap B)$$

以及：

$$\mathscr{P}(A) + \mathscr{P}(B) - \mathscr{P}(A \cap B) \leq \mathscr{P}(A) + \mathscr{P}(B)$$

练习 10. 试证明定理 (7)。

解答. 由于：

$$\begin{aligned}
\mathscr{P}(A \cap B^c) &= \mathscr{P}(A \backslash B) \\
&= \mathscr{P}(A) - \mathscr{P}(A \cap B) \\
&= \mathscr{P}(A) - \mathscr{P}(A) \cdot \mathscr{P}(B) \\
&= \mathscr{P}(A)(1 - \mathscr{P}(B)) \\
&= \mathscr{P}(A)\mathscr{P}(B^c)
\end{aligned}$$

其他同理。

2 随机变量

练习 11. 在研究中，对收入等变量取对数是非常常见的处理手段。如果 $X > 0$ 代表总体收入，那么 $\mathbb{E}(X)$ 和 $\exp[\mathbb{E}(\ln X)]$ 哪一个更大？

解答. 比较 $\mathbb{E}(X)$ 和 $\exp[\mathbb{E}(\ln X)]$ 等价于比较 $\ln[\mathbb{E}(X)]$ 和 $\mathbb{E}(\ln X)$，由于 $\ln(\cdot)$ 函数为凹函数，根据 Jensen 不等式，有 $\ln[\mathbb{E}(X)] \geq \mathbb{E}(\ln X)$，即 $\mathbb{E}(X) \geq \exp[\mathbb{E}(\ln X)]$

练习 12. 如果 r.v.$X \sim Binomial(n,p)$，求随机变量 $Y = n - X$ 的概率质量函数。

解答. $P(Y = k) = P(n - X = k) = P(X = n - k) = \begin{pmatrix} n \\ n-k \end{pmatrix} p^{n-k}(1-p)^k$

练习 13. 求对数正态分布（$Y = e^X, X \sim N(0,1)$）的概率密度函数。

解答. Y 的分布函数：

$$\begin{aligned} P(Y \le y) &= P\left(e^X \le y\right) \\ &= P(X \le \ln y) \\ &= \Phi(\ln y) \end{aligned}$$

因而其密度函数:

$$f = \frac{1}{y}\frac{1}{\sqrt{2\pi}}e^{-\frac{(\ln(y))^2}{2}}$$

练习 14. 证明:对于一个随机变量 $X \sim F_X$,随机变量 $Y = F_X(X) \sim Uniform(0,1)$。

解答. $Y \in [0,1]$，其分布函数:

$$\begin{aligned} P(Y \le y) &= P(F_X(X) \le y) \\ &= P\left(X \le F_X^{-1}(y)\right) \\ &= \int_{-\infty}^{F_X^{-1}(y)} dF_X \\ &= F_X\left(F_X^{-1}(y)\right) - F_X(-\infty) \\ &= y - 0 = y \end{aligned}$$

因而 $Y \sim Uniform(0,1)$。

练习 15. 使用任何编程语言，通过均匀分布生成 100 个 Logistic 分布（分布函数 $\frac{e^x}{1+e^x}$）的随机数，并将理论的分布函数及其经验分布函数画在一张图中。其中经验分布函数的定义为:

$$\hat{F}(x) = \frac{1}{N}\sum_{i=1}^{N} 1(X_i \le x)$$

即样本中小于 x 的样本的 s 比例。观察两者是否贴近? 生成 1000 个数据，重复以上步骤，并比较两张图的差异。

练习 16. 使用任何编程语言，通过均匀分布生成 100 个服从泊松分布（$\lambda = 1$）的随机数，并计算均值。

3 常用分布

练习 17. 求负二项分布的方差。

解答. 由于:

$$
\begin{aligned}
E\left(Z^2\right) &= \sum_{z=r}^{\infty} z^2 \cdot \begin{pmatrix} z-1 \\ r-1 \end{pmatrix} p^r (1-p)^{z-r} \\
&= \sum_{z=r}^{\infty} z^2 \cdot \frac{(z-1)!}{(r-1)!\,(z-r)!} p^r (1-p)^{z-r} \\
&= \frac{r}{p} \sum_{z=r}^{\infty} (z+1-1) \frac{z!}{r!\,(z-r)!} p^{r+1} (1-p)^{z-r} \\
&= \frac{r}{p} \sum_{z=r}^{\infty} (z+1) \frac{z!}{r!\,(z-r)!} p^{r+1} (1-p)^{z-r} \\
&\quad -\frac{r}{p} \sum_{z=r}^{\infty} \frac{z!}{r!\,(z-r)!} p^{r+1} (1-p)^{z-r} \\
&= \frac{r}{p} \sum_{z=r}^{\infty} \frac{(z+1)!}{r!\,(z-r)!} p^{r+1} (1-p)^{z-r} - \frac{r}{p} \\
&= \frac{r}{p} \sum_{z=r}^{\infty} \frac{(r+1)}{p} \frac{z'!}{r'!\,(z'-r')!} p^{r'+1} (1-p)^{z-r} - \frac{r}{p} \\
&= \frac{r(r+1)}{p^2} - \frac{r}{p} = \frac{r^2+r-rp}{p^2} \\
&\quad (r' = r+1, z' = z+1)
\end{aligned}
$$

从而 $\mathrm{Var}(Z) = \mathbb{E}\left(Z^2\right) - [\mathbb{E}(Z)]^2 = \frac{r^2+r-rp}{p^2} - \frac{r^2}{p^2} = \frac{r(1-p)}{p^2}$

练习 18. 证明几何分布的无记忆性。

解答. 根据条件概率的定义，由于 $s > t$，

$$
P(V > s | V > t) = \frac{P(V > s, V > t)}{P(V > t)} = \frac{P(V > s)}{P(V > t)} = \frac{p^s}{p^t} = p^{s-t} = P(V > s-t)
$$

练习 19. 计算泊松分布的方差。

解答. 由于:

$$
\begin{aligned}
\mathbb{E}\left(X^2\right) &= e^{-\lambda}\sum_{x=0}^{\infty} x^2 \frac{\lambda^x}{x!} \\
&= \lambda e^{-\lambda}\sum_{x=1}^{\infty} x \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda}\sum_{x=1}^{\infty} (x-1+1) \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda}\sum_{x=1}^{\infty} (x-1) \frac{\lambda^{x-1}}{(x-1)!} + \lambda e^{-\lambda}\sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda}\sum_{x'=0}^{\infty} x' \frac{\lambda^{x'}}{x'!} + \lambda \\
&= \lambda^2 + \lambda \\
&\left(x' = x-1\right)
\end{aligned}
$$

从而其方差: $\mathrm{Var}\left(X\right) = \mathbb{E}\left(X^2\right) - \left[\mathbb{E}\left(X\right)\right]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

练习 20. 计算 $\Gamma\left(\alpha, \beta\right)$ 分布的方差。

解答. 由于:

$$
\begin{aligned}
\mathbb{E}\left(X^2\right) &= \int_0^{\infty} x^2 \frac{1}{\Gamma\left(\alpha\right)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta} dx \\
&= \alpha\beta \int_0^{\infty} x \frac{1}{\Gamma\left(\alpha+1\right)\beta^{\alpha+1}} x^{\alpha} e^{-x/\beta} dx \\
&= \alpha\beta \int_0^{\infty} \frac{1}{\Gamma\left(\alpha+1\right)\beta^{\alpha+1}} x^{\alpha+1} e^{-x/\beta} dx \\
&= \alpha\beta \int_0^{\infty} \frac{1}{\Gamma\left(\alpha+2\right)\beta^{\alpha+2}} \left(\alpha+1\right)\beta x^{\alpha+1} e^{-x/\beta} dx \\
&= \alpha\beta\left(\alpha+1\right)\beta \int_0^{\infty} \frac{1}{\Gamma\left(\alpha+2\right)\beta^{\alpha+2}} x^{\alpha+1} e^{-x/\beta} dx \\
&= \alpha\left(\alpha+1\right)\beta^2
\end{aligned}
$$

因而 $\mathrm{Var}\left(X\right) = \mathbb{E}\left(X^2\right) - \left[\mathbb{E}\left(X\right)\right]^2 = \alpha\left(\alpha+1\right)\beta^2 - \alpha^2\beta^2 = \alpha\beta^2$

练习 21. 求标准正态分布的偏度、峰度。

解答. 根据标准分布的对称性，$\mathbb{E}\left(X^3\right)=0$，而

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^2 e^{-\frac{x^2}{2}}dx &= \frac{1}{\sqrt{2\pi}}\frac{1}{2}\int_{\mathbb{R}} x e^{-\frac{x^2}{2}}dx^2 \\
&= \frac{1}{\sqrt{2\pi}}\left[\frac{1}{2}x^3 e^{-\frac{x^2}{2}}\right]_{-\infty}^{\infty} - \frac{1}{2}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^2 dx e^{-\frac{x^2}{2}} \\
&= -\frac{1}{2}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^2\left(1-x^2\right)e^{-\frac{x^2}{2}}dx \\
&= -\frac{1}{2}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^2 dx + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^4 e^{-\frac{x^2}{2}}dx
\end{aligned}$$

从而：

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^4 e^{-\frac{x^2}{2}}dx &= 3\cdot\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} x^2 dx \\
&= 3
\end{aligned}$$

练习 22. Pareto 分布的密度函数为：

$$f\left(x|a,\beta\right)=\beta\cdot a^{\beta}x^{-(\beta+1)}, x>a$$

那么 Pareto 分布是否属于指数分布族?

解答. $f\left(x|a,\beta\right)=1_{(a,\infty)}\left(x\right)\exp\left\{-\left(\beta+1\right)\cdot x+\ln\beta+\beta\ln a\right\}$，由于 $1_{(a,\infty)}\left(x\right)$ 不仅仅依赖于 x 还依赖于 a，所以不属于指数分布族。

练习 23. 请用定理 (1) 计算指数分布、泊松分布的期望和方差。

解答. 对于指数分布，其密度函数：$f\left(x|\beta\right)=1_{(0,\infty)}\left(x\right)\exp\left\{-x\cdot\frac{1}{\beta}-\ln\beta\right\}$，可知 $T\left(X\right)=X$。令 $\lambda=-\frac{1}{\beta}$，那么其规范形式为：$f\left(x|\beta\right)=1_{(0,\infty)}\left(x\right)\exp\left\{x\cdot\lambda-\ln\left[-\frac{1}{\lambda}\right]\right\}$，因而

$$\mathbb{E}\left(X\right)=\frac{\partial\ln\left[-\frac{1}{\lambda}\right]}{\partial\lambda}=-\lambda\cdot\frac{1}{\lambda^2}=-\frac{1}{\lambda}=\beta$$

而其方差为：

$$\mathrm{Var}\left(X\right)=\frac{\partial^2\ln\left[-\frac{1}{\lambda}\right]}{\partial\lambda^2}=\frac{1}{\lambda^2}=\beta^2$$

对于泊松分布，其密度函数：$P\left(x|\lambda\right)=\frac{1}{x!}\exp\left\{x\ln\left(\lambda\right)-\lambda\right\}$，可知 $T\left(X\right)=X$。令 $\delta=\ln\lambda$，其规范形式为：$P\left(x|\lambda\right)=\frac{1}{x!}\exp\left\{x\delta-e^{\delta}\right\}$。从而 $\mathbb{E}\left(X\right)=e^{\delta}=\lambda$，$\mathrm{Var}\left(X\right)=e^{\delta}=\lambda$。

练习 24. 对于一个向量 $x\in\mathbb{R}^n$ 以及一个权重向量 $w\in\mathbb{R}^n$，$\sum_{i=1}^{n}w_i=\iota' w=1$，我们希望计算其加权平均：

$$\bar{x}_w=\sum_{i=1}^{n}w_i\cdot x_i$$

请写出一个幂等矩阵 P_w 使得 $P_w x=\bar{x}_w\iota$。

解答. 令 $P_\omega = \iota\omega'$，从而 $P_\omega^2 = \iota\omega'\iota\omega' = \iota\omega' = P_\omega$，所以 P_ω 为幂等矩阵。同时 $\omega' x = \bar{x}_\omega$，所以 $P_\omega x = \iota\omega' x = \bar{x}_\omega \iota$

练习 25. 若随即向量 (U, V) 的分布函数为：

$$F_{U,V}(u,v) = \frac{uv}{1-\theta(1-u)(1-v)}, \theta \in [-1,1)$$

其中 $P(U \in [0,1]) = 1, P(V \in [0,1]) = 1$，求其边缘分布函数和边缘密度函数。

解答. U 的边缘分布：

$$\begin{aligned}
F_U(u) &= F_{U,V}(u,\infty) \\
&= F_{U,V}(u,1) \\
&= \frac{u}{1-\theta(1-u)(1-1)} \\
&= u
\end{aligned}$$

因而边缘密度函数 $f_U(u) = 1, 0 \le u \le 1$。

练习 26. 如果一个随机变量 $X \sim N(0,1)$，现如下定义随机变量 Y：

$$Y = \begin{cases} X - 2 & \textit{with prob } 0.5 \\ X + 2 & \textit{with prob } 0.5 \end{cases}$$

求 $\mathrm{Var}(Y)$。

解答. 如果令随机变量 $Z \sim Ber(0.5)$，且独立于 X，那么 $Y = Z \cdot (X-2) + (1-Z) \cdot (X+2) = X + 2 - 4Z$。因而：$\mathbb{E}(Y) = \mathbb{E}(X + 2 - 4Z) = 0 + 2 - 4\mathbb{E}(Z) = 2 - 4 \cdot 0.5 = 0$，同时

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathrm{Var}(X + 2 - 4Z) \\
&= \mathrm{Var}(X) + 16\mathrm{Var}(X) \\
&= 1 + 16 \cdot 0.5 \cdot 0.5 = 5
\end{aligned}$$

练习 27. 证明 $g(X) \cdot \mathbb{E}(Y|X) = \arg\min_{h \in \mathbb{H}} \left\{ \mathbb{E}\left[(g(X) \cdot Y - h(X))^2 \right] \right\}$。

解答. 反证法，假设

$$h_0(X) = \arg\min_{h \in \mathbb{H}} \left\{ \mathbb{E}\left[(g(X) \cdot Y - h(X))^2 \right] \right\}$$

且 $h_0(X) \neq g(X) \cdot \mathbb{E}(Y|X)$，那么令

$$h'(X) = h_0(X) + \frac{\mathbb{E}\left[(g(X)Y - h_0(X))\, g(X)\, \mathbb{E}(Y|X) \right]}{\mathbb{E}\left[(g(X)\, \mathbb{E}(Y|X))^2 \right]} g(X)\, \mathbb{E}(Y|X)$$

验证 $\mathbb{E}\left[(g(X)\cdot Y-h'(X))^2\right]<\mathbb{E}\left[(g(X)\cdot Y-h_0(X))^2\right]$，因而 $h_0(X)$ 不可能是原问题的解。注意在分子上，

$$\begin{aligned}\mathbb{E}\left[(g(X)Y-h_0(X))\,g(X)\,\mathbb{E}(Y|X)\right]&=\mathbb{E}\left\{\mathbb{E}\left[(g(X)Y-h_0(X))\right]|X\right\}\\&=\mathbb{E}\left\{g(X)\,\mathbb{E}(Y|X)\,\mathbb{E}\left[(g(X)Y-h_0(X))\,|X\right]\right\}\\&=\mathbb{E}\left\{g(X)\,\mathbb{E}(Y|X)\left[g(X)\,\mathbb{E}(Y|X)-h_0(X)\right]\right\}\end{aligned}$$

练习 28. 证明 $\mathrm{Var}(Y)=\mathrm{Var}\left[\mathbb{E}(Y|X)\right]+\mathbb{E}\left[\mathrm{Var}(Y|X)\right]$

解答. 两项分开来看:

$$\begin{aligned}\mathrm{Var}\left[\mathbb{E}(Y|X)\right]&=\mathbb{E}[\mathbb{E}(Y|X)-\mathbb{E}(Y)]^2\\&=\mathbb{E}\left[(\mathbb{E}(Y|X))^2+(\mathbb{E}(Y))^2-2\mathbb{E}(Y)\,\mathbb{E}(Y|X)\right]\\&=\mathbb{E}\left[(\mathbb{E}(Y|X))^2\right]-(\mathbb{E}(Y))^2\\\mathbb{E}\left[\mathrm{Var}(Y|X)\right]&=\mathbb{E}\left(Y^2\right)-\mathbb{E}\left[[\mathbb{E}(Y|X)]^2\right]\end{aligned}$$

练习 29. 使用练习 (28) 中的结论，计算例 (??) 中的 $\mathrm{Var}(M)$。

解答. 由于 $\mathrm{Var}(M)=\mathrm{Var}(\mathbb{E}(M|N))+\mathbb{E}(\mathrm{Var}(M|N))$, 其中 $\mathbb{E}(M|N)=Np$, 因而 $\mathrm{Var}(\mathbb{E}(M|N))=\mathrm{Var}(Np)=p^2\mathrm{Var}(N)=p^2\lambda$，而 $\mathrm{Var}(M|N)=Np(1-p)$，从而 $\mathbb{E}(\mathrm{Var}(M|N))=\mathbb{E}(Np(1-p))=\lambda p(1-p)$。从而 $\mathrm{Var}(M)=p^2\lambda+\lambda p-\lambda p^2=\lambda p$。

练习 30. 如果随机变量 X 和 Y 相互独立，求 $\mathbb{E}(Y|X)$。

解答. 由于 X 和 Y 相互独立，从而:

$$\begin{aligned}\mathbb{E}(Y|X)&=\int yf_{Y|X}(y|x)\,dy\\&=\int y\frac{f_{X,Y}(x,y)}{f_X(x)}dy\\&=\int y\frac{f_X(x)\,f_Y(y)}{f_X(x)}dy\\&=\mathbb{E}(Y)\end{aligned}$$

练习 31. 使用上述结论，产生一组二维正态随机变量 $X=(X_1,X_2)$，使得第一个分量方差为 1，第二个分量方差为 2，且其相关系数为 0.5。

练习 32. Γ 分布是否属于指数分布族?

解答. 其密度函数：

$$
\begin{aligned}
f(x|\alpha,\beta) &= 1_{(0,\infty)}\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}e^{-x/\beta} \\
&= 1_{(0,\infty)}\exp\left\{(\alpha-1)\ln x-\frac{1}{\beta}x-\ln(\Gamma(\alpha))-\alpha\ln\beta\right\}
\end{aligned}
$$

因而属于指数分布族。

练习 33. 使用正态分布的规范形式求 $\mathbb{E}X^3$ 及 $\mathbb{E}X^4$，并验证 $\frac{\partial^2 C(\lambda)}{\partial\lambda\partial\lambda'}$ 的正定性。

解答. 对于正态分布，其中 $T(X)=\left(X^2,X\right)'$，且

$$
\frac{\partial C(\lambda)}{\partial\lambda}=\begin{bmatrix}\frac{\lambda_2^2}{4\lambda_1^2}-\frac{1}{2\lambda_1}\\ -\frac{\lambda_2}{2\lambda_1}\end{bmatrix}
$$

那么：

$$
\begin{aligned}
\operatorname{Var}\left(\left(X^2,X\right)'\right) &= \frac{\partial^2 C(\lambda)}{\partial\lambda\partial\lambda'} \\
&= \begin{bmatrix}\frac{1}{2\lambda_1^2}-\frac{\lambda_2^2}{2\lambda_1^3} & \frac{\lambda}{2\lambda_1^2}\\ \frac{\lambda_2}{2\lambda_1^2} & -\frac{1}{2\lambda_1}\end{bmatrix} \\
&= \begin{bmatrix}2\sigma^4+4\mu^2\sigma^2 & 2\mu\sigma^2\\ 2\mu\sigma^2 & \sigma^2\end{bmatrix} \\
&= \sigma^2\begin{bmatrix}2\sigma^2+4\mu^2 & 2\mu\\ 2\mu & 1\end{bmatrix}
\end{aligned}
$$

其正定性可以根据特征根全都大于 0 判断。

而由于：

$$
\begin{aligned}
\operatorname{Var}\left(\left(X^2,X\right)'\right) &= \mathbb{E}\left[\begin{pmatrix}X^2\\ X\end{pmatrix}\left(X^2,X\right)\right]-\mathbb{E}\left[\begin{pmatrix}X^2\\ X\end{pmatrix}\right]\mathbb{E}\left[\left(X^2,X\right)\right] \\
&= \mathbb{E}\left[\begin{pmatrix}X^2\\ X\end{pmatrix}\left(X^2,X\right)\right]-\begin{pmatrix}\mu^2+\sigma^2\\ \mu\end{pmatrix}\left(\mu^2+\sigma^2,\mu\right)
\end{aligned}
$$

从而：

$$
\begin{aligned}
\mathbb{E}\left[\begin{pmatrix}X^2\\ X\end{pmatrix}\left(X^2,X\right)\right] &= \operatorname{Var}\left(\left(X^2,X\right)'\right)+\begin{pmatrix}\mu^4+\sigma^4+2\mu^2\sigma^2 & \mu^3+\mu\sigma^2\\ \mu^3+\mu\sigma^2 & \mu^2\end{pmatrix} \\
&= \begin{pmatrix}\mu^4+3\sigma^4+6\mu^2\sigma^2 & \mu^3+3\mu\sigma^2\\ \mu^3+3\mu\sigma^2 & \mu^2+\sigma^2\end{pmatrix}
\end{aligned}
$$

练习 34. 对于一个二项分布 $X\sim Bi(N,p)$，N 已知，那么若将 p 视为变量，那

么其共轭先验是什么分布?

解答. 对于二项分布，有:

$$\begin{aligned} P(x|p) &= \begin{pmatrix} N \\ x \end{pmatrix} \exp\{x[\ln(p) - \ln(1-p)] + N\ln(1-p)\} \\ &= \begin{pmatrix} N \\ x \end{pmatrix} \exp\{x\ln(p) + (N-x)\ln(1-p)\} \end{aligned}$$

现在将参数 p 看成变量，得到:

$$\pi_t(p) = \exp\{t_1\ln(p) + t_2\ln(1-p) - \ln[k(t)]\}$$

可以看到应该是 Beta 分布

4 大样本理论

练习 35. 写程序近似逼近正态分布的分布函数并画图。

练习 36. 请找出一项表达式使其与如下序列渐进等价:

1. $\ln n + \frac{1}{2}n$
2. $\ln n + \ln(\ln n)$
3. $n^2 + e^n$

练习 37. 请确定 $a_n = \sqrt{\log n}$ 与 $b_n = \log(\sqrt{n})$ 的阶。

练习 38. 请问如下命题是否成立? 若成立，请给出证明，若不成立，请给出反例:

1. $a_n = o(b_n), c_n = o(b_n)$，那么 $a_n + c_n = o(b_n)$。
2. $a_n = o(b_n), c_n = o(d_n)$，那么 $a_n + c_n = o(b_n + d_n)$。

解答. 1. 成立，$\left|\frac{a_n+b_n}{c_n}\right| \le \left|\frac{a_n}{c_n}\right| + \left|\frac{b_n}{c_n}\right|$
2. 不成立，$a_n = \frac{1}{n^2}, b_n = \frac{1}{n}, c_n = \frac{1}{n^3}, d_n = -\frac{1}{n}$，因而 $b_n + d_n = 0$。

练习 39. 如果 $h = n^q, -1 < q < 0$，$a_n = \frac{1}{n^2h^2} + \frac{10}{n^3h}$，$b_n = 3h^3 + 10h^4$，求 q 使得 $a_n + b_n$ 以最快的速度趋向于 0。

解答. 由于 $q < 0$，因而 $\frac{n^2h^2}{n^3h} = \frac{n^q}{n} = n^{q-1} \to 0$，从而 $\frac{1}{n^3h} = o\left(\frac{1}{n^2h^2}\right)$，从而 $a_n = O\left(n^{-2}h^{-2}\right)$。而同时，$b_n = O\left(h^3\right)$，所以 $a_n + b_n = O\left(h^3 + n^{-2}h^{-2}\right)$，当 $q = \frac{2}{3}$ 时，$a_n + b_n$ 以最快的速度趋向于 0。

练习 40. 程序题：给定一个 p 和一个 n，重复的生成 n 个服从伯努利分布的随机变量，并计算其均值 $\hat{p}_n$。对于 $n = 10, 20, 30, ..., 1000$ 重复以上过程，并将结果以 n 为 x 轴，将 $\hat{p}_n$ 画在一张图上观察其收敛性。将伯努利分布换成 Cauthy 分布，再次观察其收敛性。

练习 41. 程序题：给定一个 p 和一个 n，重复的生成 n 个服从伯努利分布的随机变量，并计算 $\hat{p}_n$。对于每一个 n，重复计算出 500 个 $\hat{p}$。对于 $n = 3, 10, 30, 100$ 重复以上过程，并画出每个 n 的情况下 500 个 $\hat{p}$ 的直方图。

练习 42. 程序题：将上述练习中的伯努利分布换成正态分布的平方，重复上述练习的过程。换成 Cauthy 分布，继续重复上述练习的过程。