Homework 4

抹茶奶绿

March 13, 2025

Exercise 1.

$X \sim \mathrm{B}(6, \frac{1}{6})$，那么 $P(X=2) = \binom{6}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^4 = \frac{3125}{15552} \approx 0.2007.$

$X \overset{\text{近似}}{\sim} \mathrm{P}(1)$，那么 $P(X=2) = e^{-1}\frac{1}{2!} = \frac{1}{2e} \approx 0.1839.$

Exercise 2.

$X \sim \mathrm{B}(10^6, 2\times 10^{-6})$，
那么

$$P(X \geq 3) = 1 - P(X \leq 2) = 1 - \sum_{i=0}^{2}\binom{10^6}{i}\left(2\times 10^{-6}\right)^i\left(1-2\times 10^{-6}\right)^{10^6-i} \approx 0.3233$$

$X \overset{\text{近似}}{\sim} \mathrm{P}(2)$，
那么

$$P(X \geq 3) = 1 - P(X \leq 2) = 1 - \sum_{i=0}^{2} e^{-2}\frac{2^i}{i!} \approx 0.3233$$

Exercise 3.

记 X 为产卵个数，Y 为发育成虫的个数，那么

$$\begin{aligned} P(Y=k) &= \sum_{n=k}^{\infty} P(Y=k|X=n)P(X=n) \\ &= \sum_{n=k}^{\infty}\binom{n}{k}p^k(1-p)^{n-k}\frac{\lambda^n}{n!}e^{-\lambda} \\ &= \frac{(\lambda p)^k}{k!}e^{-\lambda p} \end{aligned}$$

那么

$$P(Y=k) = \sum_{n=k}^{\infty} P(Y=k|X=n)P(X=n) = \sum_{n=k}^{\infty}\binom{n}{k}p^k(1-p)^{n-k}\frac{\lambda^n}{n!}e^{-\lambda} = \frac{(\lambda p)^k}{k!}e^{-\lambda p}$$

故 $Y \sim \mathrm{P}(\lambda p)$.

Exercise 4.

$$\begin{cases} E(x) = \int_0^1 xf(x)\,dx = \int_0^1 x(a+bx^2)\,dx = \int_0^1 (ax+bx^3)\,dx = \left.\frac{a}{2}x^2 + \frac{b}{4}x^4\right|_0^1 = \frac{a}{2} + \frac{b}{4} = \frac{2}{3}, \\ P(\Omega) = \int_0^1 f(x)\,dx = \int_0^1 (a+bx^2)\,dx = \left. ax + \frac{b}{3}x^3\right|_0^1 = a + \frac{b}{3} = 1. \end{cases}$$

$\Rightarrow a = \frac{1}{3}, b = 1.$

Exercise 5.

由于均匀分布，易见 $P(X) = \frac{10+10}{60} = \frac{1}{3}, P(Y) = \frac{10+10}{60} = \frac{1}{3}$

Exercise 6.

由题设知 $X \sim N(270, 10^2)$，那么

$$\begin{aligned} P(X \le 240 \text{ 或} X \ge 290) &= P(X \le 240) + P(X \ge 290) \\ &= P\left(\frac{X-270}{10} \le \frac{240-270}{10}\right) + P\left(\frac{X-270}{10} \ge \frac{290-270}{10}\right) \\ &= P(Z \le -3) + P(Z \ge 2) \\ &\approx 0.00135 + 0.0228 \\ &\approx 0.02415. \end{aligned}$$

Exercise 7.

由指数分布的无记忆性知 $X \sim \text{Exp}\left(\frac{1}{3}\right)$，

那么

$$\begin{aligned} P(X > 1) &= 1 - P(X \le 1) \\ &= 1 - \left(1 - e^{-\frac{1}{3}}\right) \\ &= e^{-\frac{1}{3}} \approx 0.7165. \end{aligned}$$

由于指数分布是唯一满足无记忆性的连续分布，所以若不服从，那么由条件概率公式

$$P(X > 2.5 | X > 1.5) = \frac{1 - F(2.5)}{1 - F(1.5)}$$

Exercise 8.

由题设知 $X \sim \text{Exp}(1), Y \sim \text{Exp}(\frac{1}{2})$，

那么 $P(X > c) = e^{-c} = 1 - 0.95$，故 $c = \ln 20 \approx 2.9957.$

所以 $P(Y > c) = e^{-\frac{c}{2}} = e^{-\frac{\ln 20}{2}} = e^{\ln 20^{-\frac{1}{2}}} = 20^{-\frac{1}{2}} = 0.2236.$

Exercise 9.

由 $Exercise10$ 知，Y 的 PDF 为 $f_Y(y) = \dfrac{1}{\sqrt{2\pi}\sigma y}\exp\Big(-\dfrac{(\ln y-\mu)^2}{2\sigma^2}\Big)$.

Exercise 10.

I 记 X 的 CDF 为 $F(x) = \int_{-\infty}^{x} f(t)\,dt$

当存在 x 使得 $g(x) = y$ 时，$g(x)$ 的反函数 $g^{-1}(y)$ 存在.

此时 $Y = g(X)$ 的 CDF 为

$G(y) = P(Y \le y) = P(g(X) \le y) = P(X \le g^{-1}(y)) = F_X(g^{-1}(y)) = \int_{-\infty}^{g^{-1}(y)} f_X(t)\,dt$.

于是 Y 的 PDF 为 $f_X(g^{-1}(y)) \cdot \left|\dfrac{d}{dy}g^{-1}(y)\right|$.

当对于任意 x，$g(x) \ne y$ 时，$g^{-1}(y)$ 不存在，此时 Y 的 PDF 为 0.

II 类似的，我们有

$F(y) = P(Y \le y) = P(F(X) \le y) = P(X \le F^{-1}(y)) = F(F^{-1}(y)) = y$.

即 $Y \sim \mathrm{U}(0,1)$.

III 命 $X = F^{-1}(Y)$，那么 $F_X(x) = P(X \le x) = P(F^{-1}(Y) \le x) = P(Y \le F(x)) = F(x)$.

指数分布的 CDF 为 $F(x) = 1 - e^{-\lambda x}$，

故若 $Y \sim \mathrm{U}(0,1)$，那么 $F^{-1}(Y) = \dfrac{-\ln(1-Y)}{\lambda}$ 服从指数分布.

IV 利用概率积分变换，可以从均匀分布生成服从任意分布的随机数.($Inversetransformsampling$)

V $F(x)$ 不严格单调，那么 $F^{-1}(y)$ 不唯一，此时 $F^{-1}(y)$ 的选择会影响 Y 的 PDF，即不再是均匀分布.

Exercise 11.

显然 $P(Y = i) = p_i, i = 1, 2, \cdots, n$

对于一般的离散型随机变量 Z，也可以有同样效果，此即用均匀分布生成离散型随机变量的方法

Exercise 12.

设断开的位置为 X，那么 $X \sim \mathrm{U}(0,1)$，其 PDF 为 $f(x) = 1$，$0 \le x \le 1$.

设含固定点 $x = p_0$ 的线段长度为 Y，

那么

$$\begin{aligned}
E(Y) &= \int_0^{p_0} (1-x)\,f(x)\,dx + \int_{p_0}^{1} x\,f(x)\,dx \\
&= \int_0^{p_0} (1-x)\,dx + \int_{p_0}^{1} x\,dx \\
&= x - \frac{x^2}{2}\Big|_0^{p_0} + \frac{x^2}{2}\Big|_{p_0}^{1} \\
&= \left(p_0 - \frac{p_0^2}{2}\right) + \left(\frac{1}{2} - \frac{p_0^2}{2}\right) \\
&= p_0 - p_0^2 + \frac{1}{2}.
\end{aligned}$$

Exercise 13.

由题设知 X 的 PDF 为

$$f(x)=\begin{cases}\frac{1}{2}, & 0\le x\le 1\\ \frac{1}{2}, & 3\le x\le 4\\ 0, & \text{其他}\end{cases}$$

那么

$$E(X)=\int_0^1 x\cdot\frac{1}{2}\,dx+\int_3^4 x\cdot\frac{1}{2}\,dx=\frac{x^2}{4}\bigg|_0^1+\frac{x^2}{4}\bigg|_3^4=\frac{1}{4}+\frac{16-9}{4}=\frac{1}{4}+\frac{7}{4}=2.$$

$$E(X^2)=\int_0^1 x^2\cdot\frac{1}{2}\,dx+\int_3^4 x^2\cdot\frac{1}{2}\,dx=\frac{x^3}{6}\bigg|_0^1+\frac{x^3}{6}\bigg|_3^4=\frac{1}{6}+\frac{64-27}{6}=\frac{1}{6}+\frac{37}{6}=\frac{38}{6}=\frac{19}{3}.$$

$$\mathrm{Var}(X)=E(X^2)-[E(X)]^2=\frac{19}{3}-2^2=\frac{19}{3}-4=\frac{19}{3}-\frac{12}{3}=\frac{7}{3}.$$

Exercise 14.

设 $X\sim \mathrm{Be}(\alpha,\beta)$，

那么 X 的 PDF 为 $f(x)=\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}=\dfrac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}$.

故期望为

$$\begin{aligned}E(X)&=\int_0^1 x\,f(x)\,dx\\&=\int_0^1 x\cdot\frac{1}{B(\alpha,\beta)}\,x^{\alpha-1}(1-x)^{\beta-1}\,dx\\&=\frac{1}{B(\alpha,\beta)}\int_0^1 x^{\alpha}(1-x)^{\beta-1}\,dx\\&=\frac{B(\alpha+1,\beta)}{B(\alpha,\beta)}\\&=\frac{\alpha}{\alpha+\beta},\end{aligned}$$

下面计算方差

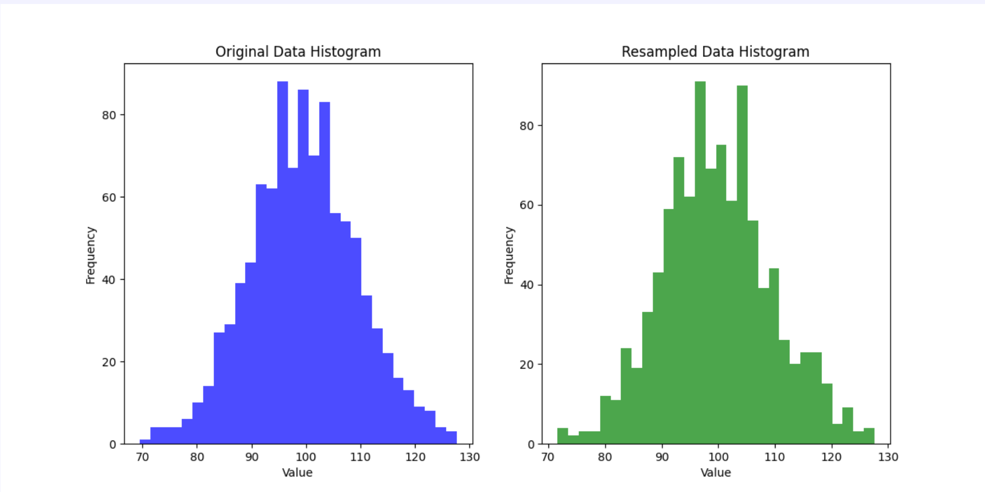
$$\begin{aligned}E(X^2)&=\frac{B(\alpha+2,\beta)}{B(\alpha,\beta)}\\&=\frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)},\end{aligned}$$

$$\begin{aligned}Var(X)&=E(X^2)-[E(X)]^2\\&=\frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}-\left(\frac{\alpha}{\alpha+\beta}\right)^2\\&=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.\end{aligned}$$

对于 $Be(1,1)$，其 PDF 为 $f(x)=1$，$0\le x\le 1$，
对于 $U(0,1)$，其 PDF 为 $f(x)=1$，$0\le x\le 1$，
二者完全一致！

Exercise 15.

```python
import numpy as np
import matplotlib.pyplot as plt

np.random.seed(0)
mu, sigma = 100, 10
data = np.random.normal(mu, sigma, 1000)
resampled_data = np.random.choice(data, 1000, replace=True)

plt.figure(figsize=(12, 6))

plt.subplot(1, 2, 1)
plt.hist(data, bins=30, color='blue', alpha=0.7)
plt.title('Original Data Histogram')
plt.xlabel('Value')
plt.ylabel('Frequency')

plt.subplot(1, 2, 2)
plt.hist(resampled_data, bins=30, color='green', alpha=0.7)
plt.title('Resampled Data Histogram')
plt.xlabel('Value')
plt.ylabel('Frequency')

plt.show()

original_mean = np.mean(data)
original_variance = np.var(data)
resampled_mean = np.mean(resampled_data)
resampled_variance = np.var(resampled_data)

print(f"Original Mean: {original_mean} Variance: {original_variance}")
print(f"Resampled Mean: {resampled_mean} Variance: {resampled_variance}")
```



$$OriginalMean : 99.54743292509804 \quad Variance : 97.42344563121542$$

$$ResampledMean : 99.7767678978011 \quad Variance : 94.19597846442375$$