第八次作业参考解答

1. 不正确.强大数定律只保证**以概率1收敛**,即$P(\lim_{n\to\infty}\frac{m}{n}=P(A))=1$,它是$\lim_{n\to\infty}\frac{m}{n}=P(A)$的必要不充分条件。

2. 记$Y_n=\sum_{i=1}^{n}(X_i-\mu_i),\forall\epsilon>0$， $\mathrm{E}[Y_n]=0,\ \mathrm{var}[Y_n]=\sum_{i=1}^{n}\sigma_i^2$
应用Chebyshev不等式:

$$\mathrm{P}\left\{\left|\frac{Y_n}{n}\right|\geq\epsilon\right\}\leq\frac{1}{n^2\epsilon^2}\mathrm{Var}[Y_n]=\frac{1}{n^2\epsilon^2}\sum_{i=1}^{n}\sigma_i^2<\frac{c}{n\epsilon^2}.$$

3. 假设$\{X_i\}_i$独立同分布且具有有限的期望μ和方差σ^2,记

$$\Phi_n(x):=\mathrm{P}\left\{\frac{\sqrt{n}}{\sigma}(\bar{X}-\mu)\leq x\right\},$$

$\forall\epsilon>0,\forall\delta>0$,可以先后证明:

I.存在足够大的$A_\delta>0$使得$\Phi(A_\delta)\geq 1-\delta$.

II.由中心极限定理$\lim_n\Phi_n(x)=\Phi(x)$知存在足够大正整数$N_1>0$使得

$$|\Phi_n(A_\delta)-\Phi(A_\delta)|<\delta,|\Phi_n(-A_\delta)-\Phi(-A_\delta)|<\delta,\forall n\geq N_1.$$

III.取足够大正整数N_2使得$\frac{\epsilon\sqrt{N_2}}{\sigma}>A_\delta$.从而当$n>\max\{N_1,N_2\}$时,

$$\begin{aligned}\mathrm{P}\left\{\left|\bar{X}-\mu\right|\geq\epsilon\right\}&\leq 1-\Phi_n(\frac{\epsilon\sqrt{n}}{\sigma})+\Phi_n(-\frac{\epsilon\sqrt{n}}{\sigma})\leq 1-\Phi_n(A_\delta)+\Phi_n(-A_\delta)\\&\leq 1-(\Phi(A_\delta)-\delta)+(\Phi(-A_\delta)+\delta)\leq 4\delta.\end{aligned}$$

即证$\lim_n\mathrm{P}\left\{\left|\bar{X}-\mu\right|\geq\epsilon\right\}=0,\forall\epsilon>0.$

4. 注意到恒等式

$$S^2-\sigma^2=\frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-\sigma^2\right]-\frac{n}{n-1}(\bar{X}-\mu)^2-\frac{\sigma^2}{n-1}.$$

应用弱大数定律知

$$\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2-\sigma^2\xrightarrow{\mathrm{P}}0,\ \bar{X}-\mu\xrightarrow{\mathrm{P}}0.$$

结合**依概率收敛**的四则运算性质[1]立即得$S^2-\sigma^2\xrightarrow{\mathrm{P}}0.$

5. 依照题目中的事件考虑，有$C\subset A\cup B$. 这是因为事件A^c表示X_n收敛于a,事件B^c表示Y_n收敛于b,而根据极限的四则运算性质$P(C^c|A^cB^c)=1$. 那么

$$0\leq P(C)\leq P(A\cup B)\leq P(A)+P(B)=0$$

从而$1-P(C)=1$,即$\frac{X_n}{Y_n}$以概率1收敛于$\frac{a}{b}$.

[1]四则运算性质指的是,若$X_n\xrightarrow{\mathrm{P}}a,Y_n\xrightarrow{\mathrm{P}}b$,则有$cX_n$,$X_n\pm Y_n$,$X_nY_n$分别依概率收敛到$ca,a\pm b,ab$.若还满足$Y_n\neq 0,b\neq 0$,则$X_n/Y_n\xrightarrow{\mathrm{P}}a/b$.证明留作练习.

6. 令$\bar{X} = \frac{X_1+X_2+\ldots X_n}{n}$，$\bar{Y} = \frac{Y_1+Y_2+\ldots Y_n}{n}$，可知这两个随机变量依概率收敛至1， $\bar{X} \xrightarrow{\text{a.s.}} 2, \bar{Y} \xrightarrow{\text{a.s.}} 5$，结合**Q5**结论可以得到**Q6**

7. 直接应用中心极限定理.

$$n\mu = 20, \sqrt{n}\sigma = \sqrt{10}, \mathrm{P}(X=20) = \mathrm{P}\left(\frac{X-20}{\sqrt{10}} = 0\right) \approx \Phi\left(\frac{0.5}{\sqrt{10}}\right) - \Phi\left(\frac{-0.5}{\sqrt{10}}\right) = 0.1256.$$

以及精确值

$$\mathrm{P}(X=20) = 2^{-40}\binom{40}{20} \approx 0.1254.$$

8. 事故数$Y = \sum\limits_{i=1}^{n} X_i \sim B(n,p)$,其中$n = 10000, p = 0.001, \mu = 10, \sigma = \sqrt{9.99}$,公司获得至少毛利润$m$千元的概率为

$$p_m = \mathrm{P}\left(20 - Y \geq m\right) = \mathrm{P}\left(\frac{Y-10}{\sqrt{9.99}} \leq \frac{10-m}{\sqrt{9.99}}\right) \approx \Phi\left(\frac{10.5-m}{\sqrt{9.99}}\right).$$

(1)公司能够获利概率$p_0 \approx \Phi\left(\frac{10.5}{\sqrt{9.99}}\right) = 0.9996$,是大概率事件,故定价合理.

(2)所求即$p_4 \approx \Phi\left(\frac{6.5}{\sqrt{9.99}}\right) = 0.9801$.

(2)所求即$\Phi\left(\frac{10.5-m}{\sqrt{9.99}}\right) = 0.95$,解得$m = 5.3011$千元.

9. $X_i \sim \mathrm{U}(-1,1), \mathrm{E}[\bar{X}] = 0, \mathrm{Var}[\bar{X}] = \dfrac{1}{3n}$,,所求$n$次测量平均误差低于$\epsilon$的概率为

$$p_{n,\epsilon} = \mathrm{P}\left(\left|\bar{X}\right| < \epsilon\right) = \mathrm{P}\left(\left|\sqrt{3n}\bar{X}\right| < \sqrt{3n}\epsilon\right) \approx 2\Phi\left(\sqrt{3n}\epsilon\right) - 1,$$

(1)$p_{25,0.2} \approx 2\Phi(\sqrt{75} \times 0.2) - 1 = 0.9167$.

(2)所求即$2\Phi(\sqrt{3n}\epsilon) - 1 \geq 1 - \alpha$,对$\epsilon = 0.2, \alpha = 0.05$解得$n \geq 33$.

(2)由Chebyshev不等式得

$$\mathrm{P}\left(\left|\bar{X}\right| < \epsilon\right) \geq 1 - \frac{\mathrm{Var}[\bar{X}]}{\epsilon^2} = 1 - \frac{1}{3n\epsilon^2},$$

对$\epsilon = 0.2, \alpha = 0.05$,令$1 - \dfrac{1}{3n\epsilon^2} \geq 1 - \alpha$解得$n \geq 167$.

10. $X \sim \mathrm{B}(5000, 0.8), \mathrm{E}[X] = 4000, \mathrm{Var}[X] = 800$,所求即

$$\mathrm{P}\left(\left|\frac{X}{5000} - 0.8\right| \leq \epsilon\right) = \mathrm{P}\left(\left|\frac{X-4000}{\sqrt{800}}\right| \leq \frac{5000\epsilon}{\sqrt{800}}\right) \approx 2\Phi\left(\frac{5000\epsilon + 0.5}{\sqrt{800}}\right) - 1,$$

令$2\Phi\left(\dfrac{5000\epsilon+0.5}{\sqrt{800}}\right) - 1 = 0.99$解得$\epsilon = 0.0145$,对应合格品数的范围$[3927, 4073]$.

11. [From 2023010455] 例如的15年全国人口抽样调查，预先假设了男、女性别比例在某地区分布是均匀的。那么在统计性男女比例时，预设了精度和置信度的前提下，无须调查完所有人口，调查宅某下限的人数后，就能以足够大的把推确保性别比例在真实值的精度范围以内，且该下限与人口总数无关。

12. 根据提示，用随机变量X_i表示第i天是上涨还是下跌。那么

$$
\begin{aligned}
Y_n &= \prod_{i=1}^{n} 1.7^{X_i} 0.5^{1-X_i} \\
\ln Y_n &= \sum_{i=1}^{n} X_i \ln 1.7 + (1 - X_i) \ln 0,5 \\
&= \sum_{i=1}^{n} [X_i \ln 3.4 + \ln 0.5]
\end{aligned}
$$

(1) 由中心极限定理，$\ln Y_n$近似服从正态分布，均值为$n(\frac{\ln 3.4}{2} + \ln 0.5) = n \ln \frac{\sqrt{3.4}}{2}$. 方差为$\frac{n}{4} \ln^2 3.4$.

(2) 利用独立性,

$$
\mathrm{E}[Y_n] = \prod_{i=1}^{n} \mathrm{E}\left[1.7^{X_i} 0.5^{1-X_i}\right] = \prod_{i=1}^{n} (1/2 * 1.7 + 1/2 * 0.5) = 1.1^n.
$$

(3) 应用强大数定律知$\bar{X} \xrightarrow{\text{a.s}} \dfrac{1}{2}$,从而可以进一步验证

$$
Y_n = \left(1.7^{\bar{X}} 0.5^{1-\bar{X}}\right)^n \xrightarrow{\text{a.s.}} 0.
$$

(4) 对于(2)的结果,当$n \to \infty$时,由于股价上涨的幅度大于下降幅度,整体而言,即从均值的角度考虑,股价在以大于1的速率迭代上涨,因此$n \longrightarrow \infty$时,$\mathrm{E}[Y_n] \to \infty$。
对于(3)而言，$1.7^{\bar{X}} 0.5^{1-\bar{X}}$是一个小于1的数，当$n \to \infty$时，接近一个公比小于1的等比数列，极限为0。

13. (1)定义$I_i(x) \begin{cases} 1, & X_i \leqslant x \\ 0, & X_i > x \end{cases}$, $F_n(x) = \sum\limits_{i=1}^{n} I_i(x)$,

$$
\mathrm{E}[F_n(x)] = F(x), \mathrm{Var}[F_n(x)] = \frac{F(x)[1 - F(x)]}{n}.
$$

(2)$\forall \epsilon > 0$,利用Chebyshev不等式

$$
\mathrm{P}\left(|F_n(x) - F(x)| \geq \epsilon\right) \leq \frac{F(x)[1 - F(x)]}{n\epsilon^2} \to 0 (n \to +\infty).
$$

14. 预期是均匀分布和正态分布对应的实验结果符合中心极限定理,Cauchy分布不符合.

作业总结：

- 切比雪夫不等式是概率论中的一个重要不等式，描述了随机变量偏离其均值的概率。其数学表达式为：

  $$P(|X - \mathbb{E}[X]| \geq k\sigma) \leq \frac{1}{k^2}$$

  其中：- X 是随机变量，- $\mathbb{E}[X]$ 是 X 的期望，- σ 是 X 的标准差，- k 是大于 0 的任意常数。

  切比雪夫不等式告诉我们，对于任意的 $k > 0$，随机变量 X 偏离其均值 $\mathbb{E}[X]$ 超过 k 倍标准差的概率不会超过 $\frac{1}{k^2}$。

  中心极限定理是概率论中的一条重要定理，它说明了独立同分布的随机变量的和（或均值）随着样本量的增大，其分布会趋近于正态分布。具体而言：

  设 $X_1, X_2, \ldots, X_n$ 为独立同分布（i.i.d.）的随机变量，均值为 μ，方差为 σ^2，则对于足够大的 n，它们的样本均值 $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ 的分布趋近于正态分布：

  $$\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$