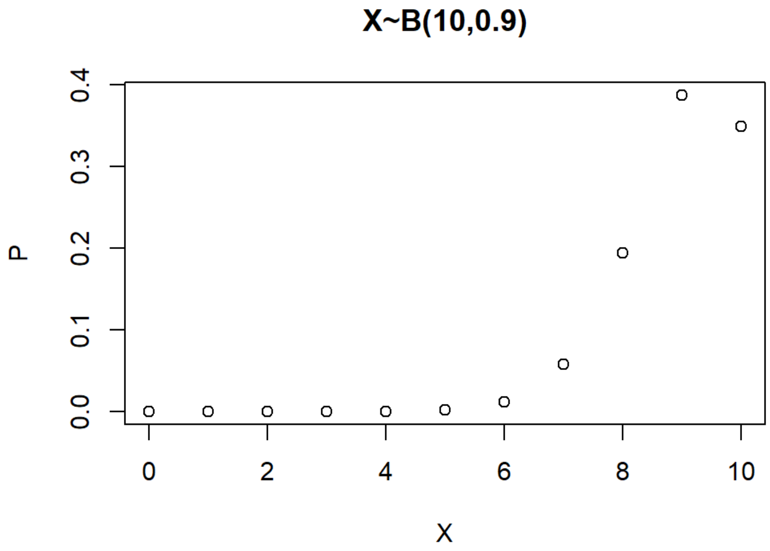
概率论与数理统计 第7次作业

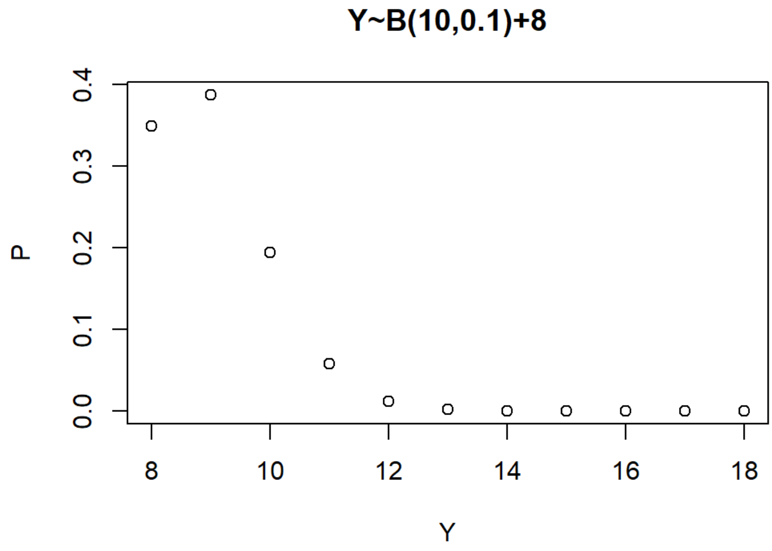
Name: 宋昊原 **Student ID:** 2022010755

November 4, 2023

1 二项分布

1.1 概率密度函数图





1.2 均值、方差、中位数、众数

均值:

$$X:9, Y:9$$

方差:

$$X:0.9, Y:0.9$$

中位数:

$$X:9, Y:9$$

众数:

$$X:9, Y:9$$

1.3 偏度系数

$$Skew(X)=-0.988,\ Skew(Y)=0.988$$

2 各种分布的矩

设$X\sim Exp(1),\ Y\sim P(4),\ Z\sim U(0,1).$

2.1

$$E(X)=1,\ Var(X)=1$$

$$Skew(X)=E((X-1)^3)=\int_0^{+\infty}(x-1)^3e^{-x}dx=-\int_0^{+\infty}(x-1)^3d(e^{-x})$$

$$=-(x-1)^3e^{-x}|_0^{+\infty}+3\int_0^{+\infty}(x-1)^2e^{-x}dx$$

$$=-1+3Var(X)=2$$

$$Kurt(X)=E((X-1)^4)=\int_0^{+\infty}(x-1)^4e^{-x}dx=-\int_0^{+\infty}(x-1)^4d(e^{-x})$$

$$=-(x-1)^4e^{-x}|_0^{+\infty}+4\int_0^{+\infty}(x-1)^3e^{-x}dx$$

$$=1+4Skew(X)=9$$

$$E(Y)=4,\ Var(Y)=4$$

$$Skew(Y)=\frac{1}{8}E((Y-4)^3)=\frac{1}{8}\sum_{k=0}^{+\infty}(k-4)^3\frac{4^ke^{-4}}{k!}$$

$$=\frac{1}{2}\sum_{k=1}^{+\infty}(k-4)^2\frac{4^{k-1}e^{-4}}{(k-1)!}-\frac{1}{2}\sum_{k=0}^{+\infty}(k-4)^2\frac{4^ke^{-4}}{k!}$$

$$=\frac{1}{2}\sum_{k=0}^{+\infty}(k-3)^2\frac{4^ke^{-4}}{k!}-\frac{1}{2}Var(Y)$$

$$=\frac{1}{2}\sum_{k=0}^{+\infty}(k-4)^2\frac{4^ke^{-4}}{k!}+4\sum_{k=1}^{+\infty}\frac{4^{k-1}e^{-4}}{(k-1)!}-\frac{7}{2}\sum_{0}^{+\infty}\frac{4^ke^{-4}}{k!}-\frac{1}{2}Var(Y)$$

$$= \frac{1}{2}Var(Y) + 4 \times 1 - \frac{7}{2} \times 1 - \frac{1}{2}Var(Y) = \frac{1}{2}$$

$$Kurt(Y) = \frac{1}{16}E((Y-4)^4) = \frac{1}{16}\sum_{k=0}^{+\infty}(k-4)^4\frac{4^k e^{-4}}{k!}$$

$$= \frac{1}{4}\sum_{k=1}^{+\infty}(k-4)^3\frac{4^{k-1}e^{-4}}{(k-1)!} - \frac{1}{4}\sum_{k=0}^{+\infty}(k-4)^3\frac{4^k e^{-4}}{k!}$$

$$= \frac{1}{4}\sum_{k=0}^{+\infty}(k-3)^3\frac{4^k e^{-4}}{k!} - 2Skew(Y)$$

$$= \frac{1}{4}\sum_{k=0}^{+\infty}((k-4)^3 + 3(k-4)^2 + 3(k-4) + 1)\frac{4^k e^{-4}}{k!} - 2Skew(Y)$$

$$= 2Skew(Y) + \frac{3}{4}Var(Y) + 0 + 1 - 2Skew(Y) = \frac{13}{4}$$

$$E(Z) = \frac{1}{2},\ Var(Y) = \frac{1}{12}$$

$$Skew(Y) = 24\sqrt{3}E((Z-\frac{1}{2})^3) = 24\sqrt{3}\int_0^1 (z-\frac{1}{2})^3 dz$$

$$= 24\sqrt{3}\int_0^1 (z^3 - \frac{3}{2}z^2 + \frac{3}{4}z - \frac{1}{8})dz$$

$$= 24\sqrt{3}(\frac{1}{4} - \frac{1}{2} + \frac{3}{8} - \frac{1}{8}) = 0$$

$$Kurt(Y) = 144E((Z-\frac{1}{2})^4) = 144\int_0^1 (z-\frac{1}{2})^4 dz$$

$$= 144\int_0^1 (z^4 - 2z^3 + \frac{3}{2}z^2 - \frac{1}{2}z + \frac{1}{16})dz$$

$$= 144(\frac{1}{5} - \frac{1}{2} + \frac{1}{2} - \frac{1}{4} + \frac{1}{16})$$

$$= \frac{9}{5}$$

2.2

$$M_X(t) = E(e^{tX}) = \int_0^{+\infty} e^{xt}e^{-x}dx$$

$$= \int_0^{+\infty} e^{(t-1)x}dx = \frac{1}{t-1}\int_0^{+\infty} e^{(t-1)x}d((t-1)x)$$

$$= \frac{1}{1-t},\ t < 1$$

$$M_Y(t) = E(e^{tY}) = \sum_{k=0}^{+\infty} e^{kt}\frac{4^k e^{-4}}{k!}$$

$$= \frac{1}{e^4}\sum_{k=0}^{+\infty}\frac{(4e^t)^k}{k!} = e^{4(e^t-1)}$$

$$M_Z(t)=E(e^{tZ})=\int_0^1 e^{zt}dz=\frac{e^t-1}{t}$$

特别地，$M_Z(0)=1$.

2.3

$$M_X'(t)=\frac{1}{(1-t)^2}\Rightarrow E(X)=M_X'(0)=1$$
$$M_X^{(2)}(t)=\frac{2}{(1-t)^3}\Rightarrow E(X^2)=M_X^{(2)}(0)=2$$
$$M_X^{(3)}(t)=\frac{6}{(1-t)^4}\Rightarrow E(X^3)=M_X^{(3)}(0)=6$$
$$M_X^{(4)}(t)=\frac{24}{(1-t)^5}\Rightarrow E(X^4)=M_X^{(4)}(0)=24$$

故

$$Skew(X)=E((X-1)^3)=E(X^3)-3E(X^2)+3E(X)-1=2$$
$$Kurt(X)=E((X-1)^4)=E(X^4)-4E(X^3)+6E(X^2)-4E(X)+1=9$$

$$M_Y'(t)=4e^te^{4(e^t-1)}\Rightarrow E(X)=M_Y'(0)=4$$
$$M_Y^{(2)}(t)=(4e^t+16e^{2t})e^{4(e^t-1)}\Rightarrow E(X^2)=M_Y^{(2)}(0)=20$$
$$M_Y^{(3)}(t)=(4e^t+48e^{2t}+64e^{3t})e^{4(e^t-1)}\Rightarrow E(X^3)=M_Y^{(3)}(0)=116$$
$$M_Y^{(4)}(t)=(4e^t+112e^{2t}+384e^{3t}+256e^{4t})e^{4(e^t-1)}\Rightarrow E(X^4)=M_Y^{(4)}(0)=756$$

故

$$E((Y-4)^3)=E(Y^3)-12E(Y^2)+48E(Y)-64=4$$
$$Skew(Y)=\frac{1}{8}E((Y-4)^3)=\frac{1}{2}$$
$$E((Y-4)^4)=E(Y^4)-16E(Y^3)+96E(Y^2)-256E(Y)+256=52$$
$$Kurt(Y)=\frac{1}{16}E((Y-4)^4)=\frac{13}{4}$$

$$M_Z(t)=\frac{e^t-1}{t}=\frac{\sum\limits_{n=1}^{+\infty}\frac{t^n}{n!}}{t}=\sum_{n=0}^{+\infty}\frac{t^n}{(n+1)!}$$

于是

$$E(X^k)=M_Z^{(k)}(0)=\frac{k!}{(k+1)!}=\frac{1}{k+1}$$

故

$$Skew(Z)=24\sqrt{3}E((X-\frac{1}{2})^3)=24\sqrt{3}(E(X^3)-\frac{3}{2}E(X^2)+\frac{3}{4}E(X)-\frac{1}{8})$$
$$=24\sqrt{3}(\frac{1}{4}-\frac{1}{2}+\frac{3}{8}-\frac{1}{8})=0$$
$$Kurt(Z)=144E((X-\frac{1}{2})^4)=144(E(X^4)-2E(X^3)+\frac{3}{2}E(X^2)-\frac{1}{2}E(X)+\frac{1}{16})$$
$$=144(\frac{1}{5}-\frac{1}{2}+\frac{1}{2}-\frac{1}{4}+\frac{1}{16})=\frac{9}{5}$$

3 由矩母函数求分布

首先考虑指数分布$Y \sim Exp(\lambda)$的矩母函数.

$$M_Y(t) = E(e^{tY}) = \int_0^{+\infty} \lambda e^{(t-\lambda)y} dy = \frac{\lambda}{\lambda - t}$$

令$Y_2 \sim Exp(2)$，$Y_3 \sim Exp(3)$，则

$$M_X(t) = \frac{1}{3} M_{Y_2}(t) + \frac{2}{3} M_{Y_3}(t)$$

故

$$f_X(x) = \frac{1}{3} f_{Y_2}(x) + \frac{2}{3} f_{Y_3}(x) = \frac{2}{3} e^{-2x} + 2e^{-3x},\ x > 0$$

4 对数正态分布

4.1

$$E(e^{tY}) = E(e^{te^X}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \exp(-\frac{(x-\mu)^2}{2\sigma^2} + te^x) dx$$

被积函数在x趋于某一侧无穷时极限为$+\infty$，故积分不存在.

4.2

$$M_X(t) = e^{\mu t + \frac{(\sigma t)^2}{2}}$$

而

$$E(Y^n) = E(e^{nX}) = M_X(n) = e^{\mu n + \frac{(\sigma n)^2}{2}}$$

5 独立Poisson分布的和

考虑X_i的矩母函数

$$M_{X_i}(t) = E(e^{tX_i}) = \sum_{k=0}^{\infty} e^{kt} \frac{\lambda_i^k e^{-\lambda_i}}{k!} = e^{-\lambda_i} \sum_{k=0}^{\infty} \frac{(\lambda_i e^t)^k}{k!} = e^{\lambda_i (e^t - 1)}$$

则

$$M_Y(t) = M_{X_1}(t) M_{X_2}(t) = e^{\lambda_1 (e^t-1) + \lambda_2 (e^t - 1)} = e^{(\lambda_1 + \lambda_2)(e^t - 1)}$$

这与$P(\lambda_1 + \lambda_2)$具有相同矩母函数，故

$$Y \sim P(\lambda_1 + \lambda_2)$$

6 两次切线段

设第一次切断后余下部分长度为X，第二次余下部分长度为Y，则

$$E(Y) = E(E(Y|X))$$

$X = x$时

$$E(Y|x) = \frac{1}{2} x$$

故

$$E(Y) = E(E(Y|X)) = E(\frac{1}{2} X) = \frac{1}{2} E(X) = \frac{1}{4}$$

7 矿工迷路

设矿工当前选择第X扇门，出去需要Y小时.
此问题是没有记忆性的，故若矿工回到原地后再次选择第X扇门，其出去仍需要的时间期望是不变的，故

$$E(Y|1) = 2,\ E(Y|2) = 3 + E(Y),\ E(Y|3) = 1 + E(Y)$$

由于

$$E(Y) = \frac{1}{3}(E(Y|1) + E(Y|2) + E(Y|3))$$

可知

$$E(Y) = \frac{1}{3}(6 + 2E(Y))$$

解得

$$E(Y) = 6$$

即矿工能走出去的平均用时是6小时.

8 协方差与条件期望

由于$E(Y|X) = X$，我们有

$$E(Y) = E(E(Y|X)) = E(X)$$

故

$$Cov(X,Y) = E((X-\mu_1)(Y-\mu_2)) = E(XY) - E(X)E(Y) = E(XY) - E^2(X)$$

其中

$$E(XY) = E(E(XY|X)) = E(XE(Y|X)) = E(X^2)$$

故

$$Cov(X,Y) = E(X^2) - E^2(X) = Var(X)$$

9 条件期望与独立性

9.1

若X, Y独立，我们有

$$E(Y|X) = \sum_y y f_{Y|X}(y|x) = \sum_y y \frac{f(x,y)}{f_X(x)} = \sum_y y f_Y(y) = E(Y)$$

这里的$\sum\limits_y$符号根据离散或连续的实际情况分别解读为求和或积分.

9.2

反之不成立.
考虑二元正态分布$(X,Y) \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$，当$\rho \neq 0$时，$X, Y$不独立. 但无论$\rho$取值如何，都满足$E(Y|X) = \mu_2 = E(Y)$.

10 方差的相对性

等式右侧

$$Var(\hat{Y}) + Var(\tilde{Y})$$

$$= E(E^2(Y|X)) - E^2(E(Y|X)) + E(Y^2) - 2E(YE(Y|X)) + E(E^2(Y|X)) - E^2(Y) + 2E(Y)E(E(Y|X)) - E^2(E(Y|X))$$

$$= Var(Y) + 2E(E^2(Y|X)) - 2E(YE(Y|X))$$

而
$$E(YE(Y|X)) = E(E(YE(Y|X)|X)) = E(E(Y|X)E(Y|X)) = E(E^2(Y|X))$$
故
$$Var(Y) = Var(\hat{Y}) + Var(\tilde{Y})$$
直观解释：Y的方差由两部分产生，一部分X改变时Y的波动，即$Var(\hat{Y})$，另一部分是X不变时Y视为一元分布的波动，即$Var(\tilde{Y})$.
这类似于物理学相对运动中的绝对速度等于相对速度加牵连速度.

11 条件方差

11.1

$$Var(Y|X) = E((Y - E(Y|X))^2|X) = E(Y^2|X) - 2E(YE(Y|X)|X) + E(E^2(Y|X)|X)$$
以X为条件时，$E(Y|X)$对每个确定是X是常数，于是
$$Var(Y|X) = E(Y^2|X) - 2E(Y|X)E(Y|X) + E^2(Y|X) = E(Y^2|X) - E^2(Y|X)$$

11.2

$$\begin{aligned} E(Var(Y|X)) + Var(E(Y|X)) &= E(E(Y^2|X)) - E(E^2(Y|X)) + E(E^2(Y|X)) - E^2(E(Y|X)) \\ &= E(E(Y^2|X)) - E^2(E(Y|X)) = E(Y^2) - E^2(Y) = Var(Y) \end{aligned}$$

12 最优预测1

$$Y_{opt} = E(Y|X = 0.5) = \frac{1}{2}\sqrt{1 - 0.5^2} = \frac{\sqrt{3}}{8}$$

13 最优预测2

13.1

$$\begin{aligned} E((Y - (aX + b))^2) &= E(Y^2) + a^2E(X^2) + b^2 - 2aE(XY) - 2bE(Y) + 2abE(X) \\ &= (\sigma_2^2 + \mu_2^2) + a^2(\sigma_1^2 + \mu_1^2) + b^2 - 2a(\rho\sigma_1\sigma_2 + \mu_1\mu_2) - 2b\mu_2 + 2ab\mu_1 \end{aligned}$$
令
$$\mathbf{A} = \begin{bmatrix} \sigma_1^2 + \mu_1^2 & \mu_1 & -(\rho\sigma_1\sigma_2 + \mu_1\mu_2) \\ \mu_1 & 1 & -\mu_2 \\ -(\rho\sigma_1\sigma_2 + \mu_1\mu_2) & -\mu_2 & \sigma_2^2 + \mu_2^2 \end{bmatrix}$$
$$\mathbf{x} = \begin{bmatrix} a \\ b \\ 1 \end{bmatrix}$$
则
$$E((Y - (aX + b))^2) = \mathbf{x}^T\mathbf{A}\mathbf{x}$$
作Cholesky分解，有
$$\mathbf{A} = \mathbf{U}^T\mathbf{U}$$
其中
$$\mathbf{U} = \begin{bmatrix} \sqrt{\sigma_1^2 + \mu_1^2} & \frac{\mu_1}{\sqrt{\sigma_1^2+\mu_1^2}} & -\frac{\rho\sigma_1\sigma_2+\mu_1\mu_2}{\sqrt{\sigma_1^2+\mu_1^2}} \\ & \frac{\sigma_1}{\sqrt{\sigma_1^2+\mu_1^2}} & \frac{\rho\mu_1\sigma_2-\mu_2\sigma_1}{\sqrt{\sigma_1^2+\mu_1^2}} \\ & & \sigma_2\sqrt{1-\rho^2} \end{bmatrix}$$
于是
$$E((Y - (aX + b))^2) = \|\mathbf{U}\mathbf{x}\|^2$$

故要使这个二次型最小，只需让$\mathbf{Ux}$的前两个分量为0，解得

$$a = \rho\frac{\sigma_2}{\sigma_1}$$

$$b = \mu_2 - \rho\frac{\sigma_2}{\sigma_1}\mu_1$$

13.2

此时的均方误差是

$$E((Y-(aX+b))^2) = \sigma_2^2(1-\rho^2)$$

当σ_2很小，即Y波动较小，或$|\rho|$很接近1，即X和Y接近线性相关时，均方误差的值接近0.

13.3

此时的最优预测是

$$\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X-\mu_1)$$

考虑

$$E(Y|x) = \int_{-\infty}^{+\infty} y f_{Y|X}(y|x)dy$$

$$= \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}\sigma_2}\int_{-\infty}^{+\infty} y\exp\left(-\frac{(y-(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1))^2)}{2(1-\rho^2)\sigma_2^2}\right)dy$$

令

$$\alpha = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$$

$$\beta = 2(1-\rho^2)\sigma_2^2$$

则

$$E(Y|x) = \frac{1}{\sqrt{\pi\beta}}\int_{-\infty}^{+\infty} y e^{-\frac{(y-\alpha)^2}{\beta}}dy$$

$$= \frac{1}{\sqrt{\pi\beta}}\left(\int_{-\infty}^{+\infty}(y-\alpha)e^{-\frac{(y-\alpha)^2}{\beta}}d(y-\alpha) + \int_{-\infty}^{+\infty}\alpha e^{-\frac{(y-\alpha)^2}{\beta}}dy\right)$$

$$= \frac{1}{\sqrt{\pi\beta}}(0+\alpha\sqrt{\pi\beta}) = \alpha = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$$

这说明

$$E(Y|X) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X-\mu_1)$$

即$E(Y|X)$就是最优线性预测.

14 随机数量的随机变量

14.1

$$E(Y) = E(E(Y|N)) = E(N\mu) = \mu E(N) = \mu a$$

14.2

$$M_N(t) = E(e^{tN}) = \sum_{n=1}^{\infty} p_n e^{nt}$$

14.3

$$M_Y(t) = E(e^{tY}) = E(E(e^{tY}|N)) = E((M_X(t))^N) = \sum_{n=1}^{\infty} p_n (M_X(t))^n = M_N(\ln M_X(t))$$

14.4

$$M_N(t)=\sum_{n=1}^{\infty}(1-p)^{n-1}pe^{nt}=\frac{pe^t}{1-(1-p)e^t},\ t<\ln\frac{1}{1-p}$$

$$M_X(t)=\frac{\lambda}{\lambda-t}$$

故

$$M_Y(t)=M_N(\ln M_X(t))=M_N(\ln\frac{\lambda}{\lambda-t})=\frac{p\lambda}{p\lambda-t}$$

故

$$Y\sim Exp(p\lambda)$$

14.5

不相同.

$$M_{X_1+X_2}(t)=(M_X(t))^2=\frac{\lambda^2}{(\lambda-t)^2}$$

15 个数随机的正态分布不再是正态分布

设相互独立的$X_i\sim N(0,1)$，则其公共的矩母函数

$$M_X(t)=e^{\frac{1}{2}t^2}$$

此时设$Y=X_1+...+X_N$，N服从与上题相同的几何分布，则

$$M_Y(t)=M_N(\frac{1}{2}t^2)=\frac{pe^{\frac{1}{2}t^2}}{1-(1-p)e^{\frac{1}{2}t^2}}$$

这不可能是正态分布.

16 计算机实验：随机徘徊

16.1

$$E(Y_i)=0$$

故

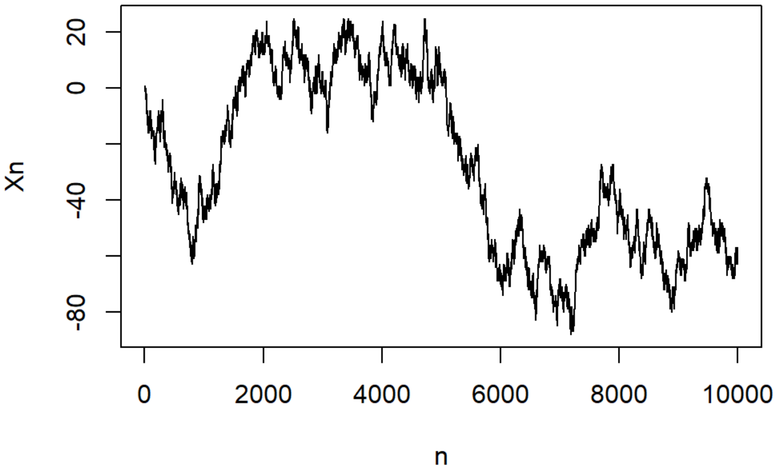
$$E(X_n)=nE(Y)=0$$

$$Var(X_n)=E(X_n^2)-E^2(X_n)=nE(Y^2)+n(n-1)E^2(Y)=\frac{n}{4}$$

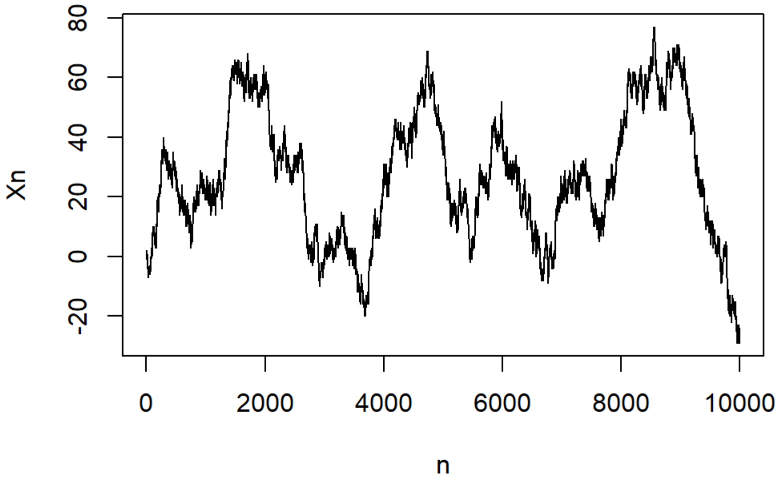
16.2

我重复了5次模拟，以下是5次的随机序列图像:

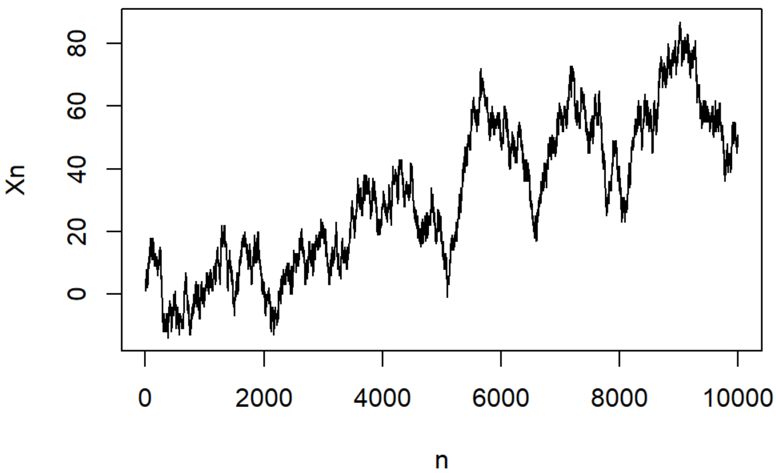
第一次


Xn
20
0
-40
-80
0
2000
4000
6000
8000
10000
n

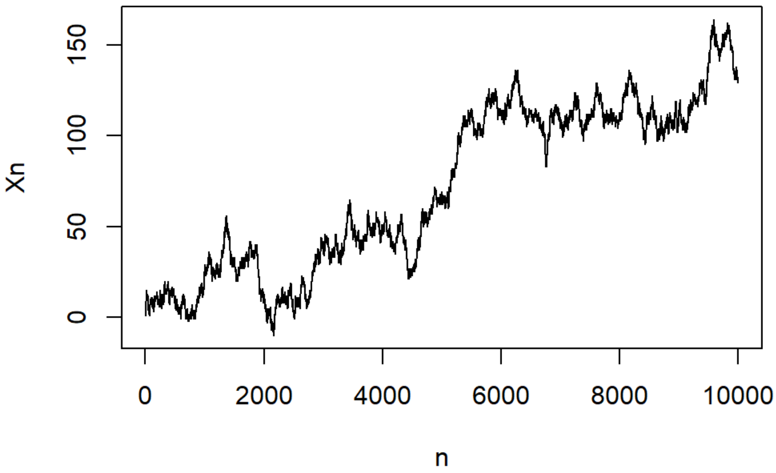
第二次


Xn
80
60
40
20
0
-20
0
2000
4000
6000
8000
10000
n

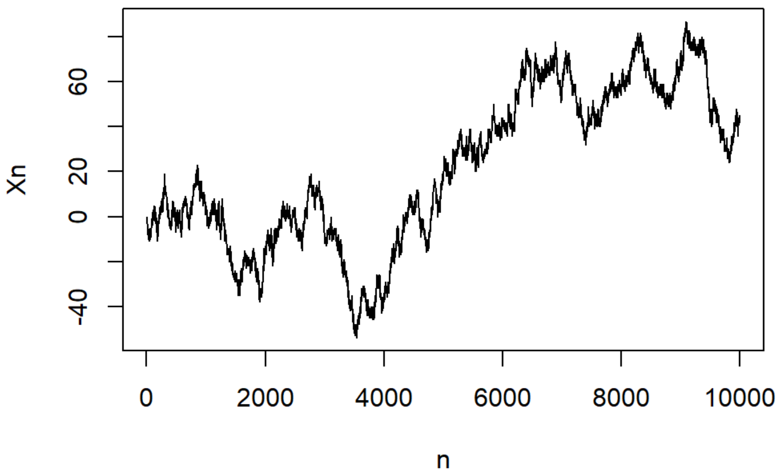
第三次


Xn
80
60
40
20
0
0
2000
4000
6000
8000
10000
n

第四次



第五次



可以看出，随机序列整体呈现出非常随机的情况，并没有明显的集中趋势，并且某种程度上n越大，X_n偏离原点的趋势越明显.
同时，不同模拟之间图形呈现出显著的区别，有的正偏有的负偏，有的（如第四次）甚至偏移出非常远.
由于$E(X_n)$，图形正偏和负偏自然是都有可能的，但因为$Var(X_n)$与n成正比，随机徘徊位置和原点的平均偏差近似与$\sqrt{n}$成正比，因此产生出n越大越不可预测的结果.
由于$Var(X_n) = \frac{n}{4}$，可以计算出X_{10000}的标准差为50，所以徘徊10000个时间周期后距离原点的距离基本在几十的数量级，较少数（如上面第四次）会偏差超过一百.