

概率论与数理统计 第3次作业

Name: 宋昊原 Student ID: 2022010755

October 9, 2023

1 随机变量的例子

1.1

有3白2黑共5个球，从中不放回地摸出三个，其中白球的个数为随机变量 X_1 .

X_1 是离散随机变量

将5个球分别标记为1,2,3,4,5，其中1至3为白球，4至5为黑球，则对应的样本空间为

$$\Omega_1 = \{(i, j, k) | i, j, k \in \{1, 2, 3, 4, 5\} \wedge i \neq j \wedge j \neq k \wedge i \neq k\}$$

1.2

有一个均匀硬币，连续掷无数次，第一次发生正反面改变（即在此之前的所有试验结果全部相同）的次数为随机变量 X_2 .

X_2 是离散随机变量

对应样本空间 Ω_2 为所有可数长度的“正”“反”串组成的集合.

1.3

在区间(0,1]中随机选取一个实数，所得为 X_3 . X_3 满足下列累积分布函数:

$$F(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x < 1 \\ 1 & x \geq 1 \end{cases} \quad (1)$$

X_3 为连续随机变量. 对应样本空间为区间(0,1].

1.4

掷五个骰子，规定若没有3个及以上相同的点数，则 $X_4 = 0$ ，否则 X_4 等于点数之和.

X_4 为离散随机变量. 样本空间 Ω_4 为长度为5的{1,2,3,4,5,6}组成的串.

1.5

任取一成年人，其身高构成随机变量 X_5 .

X_5 可以认为是连续随机变量，样本空间 Ω_5 为世界上所有成年人的集合.

2 分布函数的性质

本题需要利用概率的连续性. 故先证明之.

引理1

设 A_1, A_2, \dots 为单调递增的事件序列，即 $A_n \subset A_{n+1} (n = 1, 2, \dots)$ ，令 $A = \bigcup_{n=1}^{\infty} A_n$ ，则 $P(A) = \lim_{n \rightarrow \infty} P(A_n)$.

引理1证明

首先, 令 $B_1 = A_1$, $B_n = A_n - A_{n-1} (\forall n \geq 2)$, 则所有 $B_n (n \geq 1)$ 两两互斥, 且

$$\forall m, \sum_{n=1}^m A_n = A_m = \sum_{n=1}^m B_n$$

两边对 $m \rightarrow \infty$ 取极限, 有

$$A = \sum_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty} B_n$$

于是

$$P(A) = P\left(\sum_{n=1}^{\infty} B_n\right)$$

由 B_n 的互斥性

$$\begin{aligned} P(A) &= \sum_{n=1}^{\infty} P(B_n) \\ &= \lim_{m \rightarrow \infty} \sum_{n=1}^m P(B_n) \\ &= \lim_{m \rightarrow \infty} P\left(\sum_{n=1}^m B_m\right) \\ &= \lim_{m \rightarrow \infty} P(A_m) \end{aligned}$$

证毕.

引理2

设 A_1, A_2, \dots 为单调递减的事件序列, 即 $A_{n+1} \subset A_n (n = 1, 2, \dots)$, 令 $A = \prod_{n=1}^{\infty} A_n$, 则 $P(A) = \lim_{n \rightarrow \infty} P(A_n)$.

引理2证明

令 $B_n = A_n^c, \forall n \in \mathbb{N}$, 则

$$B_n \subset B_{n+1}, \forall n \in \mathbb{N}$$

$$\text{令 } B = \sum_{n=1}^{\infty} B_n$$

$\forall \omega \in A, \forall n \in \mathbb{N}, \omega \in A_n$, 故 $\forall n \in \mathbb{N}, \omega \notin B_n$, 故 $\omega \notin B$.

反之, $\forall \omega \notin A, \exists n_0 \in \mathbb{N}, s.t. \omega \notin A_{n_0}$, 故 $\omega \in B_{n_0}$, 故 $\omega \in B$.

于是, $B = A^c$.

由引理1

$$P(B) = \lim_{n \rightarrow \infty} P(B_n)$$

两边同时用1减, 则有

$$P(A) = \lim_{n \rightarrow \infty} P(A_n)$$

证毕.

2.1

定义事件序列

$$A_n = \{\omega \in \Omega | X(\omega) \leq n\}$$

则

$$\forall n \geq 1, A_n \subset A_{n+1}$$

即 A_n 满足引理条件, 接下来考虑 $A = \sum_{n=1}^{\infty} A_n$:

$\forall \omega \in \Omega, \omega \in A_{\lceil X(\omega) \rceil}$, 于是 $\omega \in A$.

这说明 $\Omega \subset A$, 故 $A = \Omega$.

由引理1

$$\lim_{n \rightarrow \infty} P(X \leq n) = P(A) = P(\Omega) = 1$$

于是

$$\lim_{n \rightarrow \infty} F(n) = 1$$

接下来考虑 $\lim_{x \rightarrow +\infty} F(x)$.

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, s.t. \forall n > N \wedge n \in \mathbb{N}, |F(n) - 1| < \epsilon$$

于是, 取 $M = N + 1$, 则 $\forall x > M \wedge x \in \mathbb{R}$, 由 $F(x)$ 单调性, 有

$$F(\lfloor x \rfloor) \leq F(x) \leq 1$$

而 $\lfloor x \rfloor \geq M > N$, 于是

$$|F(\lfloor x \rfloor) - 1| < \epsilon$$

故

$$|F(x) - 1| < \epsilon$$

这说明

$$\lim_{x \rightarrow +\infty} F(x) = 1$$

负无穷处的极限同理可证.

2.2

$\forall x_0 \in \mathbb{R}$, 定义事件序列

$$A_n = \{\omega \in \Omega | X(\omega) \leq x_0 + \frac{1}{n}\}$$

令 $A = \{\omega \in \Omega | X(\omega) \leq x_0\}$, 显然 $A \subset A_n (\forall n \in \mathbb{N}_+)$, 故 $A \subset \prod_{n=1}^{\infty} A_n$.

假设 $\exists \omega \in \prod_{n=1}^{\infty} A_n \wedge \omega \notin A$, 则 $X(\omega) > x_0$, 取 $n_0 = \lceil \frac{1}{X(\omega) - x_0} \rceil$, 则有 $\omega \notin A_{n_0}$, 这与假设矛盾. 即上述 ω 不存在,

即 $\prod_{n=1}^{\infty} A_n \subset A$.

故 $A = \prod_{n=1}^{\infty} A_n$. 由引理2,

$$P(A) = \lim_{n \rightarrow \infty} P(A_n)$$

故

$$F(x_0) = \lim_{n \rightarrow \infty} F(x_0 + \frac{1}{n})$$

接下来考虑 $\lim_{x \rightarrow x_0^+} F(x)$.

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, s.t. \forall n > N \wedge n \in \mathbb{N}, |F(x_0 + \frac{1}{n}) - F(x_0)| < \epsilon$$

于是, 取 $\delta = \frac{1}{N+1}$, 则 $\forall 0 \leq x - x_0 < \delta \wedge x \in \mathbb{R}$, 由 $F(x)$ 单调性, 有

$$F(x_0 + \frac{1}{\lfloor \frac{1}{x-x_0} \rfloor}) \geq F(x) \geq F(x_0)$$

而 $\lfloor \frac{1}{x-x_0} \rfloor \geq N+1 > N$, 于是

$$|F(x_0 + \frac{1}{\lfloor \frac{1}{x-x_0} \rfloor}) - F(x_0)| < \epsilon$$

故

$$|F(x) - F(x_0)| < \epsilon$$

这说明

$$\lim_{x \rightarrow x_0^+} F(x) = F(x_0)$$

即 $F(x)$ 在任意 $x_0 \in \mathbb{R}$ 右侧连续.

2.3

首先证明下述结论:

$$\lim_{x \rightarrow a^-} F(x) = P(X < a)$$

这仍然可以利用前两问使用的 $\epsilon - \delta$ 语言论述, 这里不再赘述, 只给出事件序列的构造:

$$A_n = \{\omega \in \Omega | X(\omega) \leq a - \frac{1}{n}\}$$

$$A = \{\omega \in \Omega | X(\omega) < a\}$$

可以类似地证明, 此序列满足引理1条件, 故

$$P(x < a) = \lim_{n \rightarrow \infty} F(a - \frac{1}{n})$$

利用 $F(x)$ 单调性即可类似地得到

$$P(X < a) = \lim_{x \rightarrow a^-} F(x)$$

于是

$$P(a \leq X \leq b) = P(X \leq b) - P(X < a) = F(b) - \lim_{x \rightarrow a^-} F(x) = F(b) - F(a^-)$$

3 同分布

3.1

$$P(\omega_1) = P(X = 1) = P(Y = 2) = \frac{1}{3}$$

$$P(\omega_2) = P(X = 2) = P(Y = 3) = \frac{1}{3}$$

$$P(\omega_3) = P(X = 3) = P(Y = 1) = \frac{1}{3}$$

于是, X 和 Y 均满足对 $\{1, 2, 3\}$ 中任何一个取值, 概率都为 $\frac{1}{3}$, 且没有其他可能取值. X 和 Y 同分布.

3.2

列出 $X + Y$ 、 $Y - X$ 的函数关系:

随机变量	ω_1	ω_2	ω_3
$X+Y$	3	5	4
$Y-X$	1	1	-2

于是二者的分布表:

X	3	4	5
P	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
Y	-2	1	
P	$\frac{1}{3}$	$\frac{2}{3}$	

4 方差

4.1 证明

$$\begin{aligned} \text{Var}(X) &= \sum_i (x_i - E(X))^2 p_i \\ &= \sum_i (x_i^2 p_i - 2x_i E(X) p_i + E^2(X) p_i) \\ &= \sum_i x_i^2 p_i - 2E(X) \sum_i x_i p_i + E^2(X) \sum_i p_i \\ &= E(X^2) - 2E(X) \times E(X) + E^2(X) \times 1 \\ &= E(X^2) - E^2(X) \end{aligned}$$

4.2 理解

中学学过两种方差：数据的方差和随机变量的方差，其中随机变量的方差与课上的定义一致，数据的方差主要差别在于统计一组数据时可以重复，但各个数据的权重均等。然而，这和随机变量的方差某种程度上还是一回事：

$$\begin{aligned} S^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \\ &= \frac{1}{n} \sum_j n_j (x_j - \bar{x})^2 \\ &= \sum_j \frac{n_j}{n} (x_j - \bar{x})^2 \\ &= \sum_j p_j (x_j - \bar{x})^2 \end{aligned}$$

这表示，将一组数据的统计方差中相同的元素合并，再将每个元素出现的频率视为选中之的概率，则化为随机变量的方差。

5 取球问题

5.1 不放回版

假设取球不会停止，设 A_i 为第 i 个取出白球，则

$$\begin{aligned} P(X = k) &= P(A_k | A_1^c A_2^c \dots A_{k-1}^c) P(A_1^c A_2^c \dots A_{k-1}^c) \\ &= \frac{b}{a+b-k+1} \times \frac{a(a-1)\dots(a-k+2)}{(a+b)(a+b-1)\dots(a+b-k+2)} \\ &= \frac{ba!(a+b-k)!}{(a-k+1)!(a+b)!}, \forall k = 1, 2, \dots, a+1 \end{aligned}$$

5.2 放回版

第一次取出白球出现在第 n 次意味着前面 $n-1$ 次全部取到黑球且第 n 次取到白球，这 n 次互相独立，故

$$P(X = n) = \frac{b}{a+b} \times \left(\frac{a}{a+b}\right)^{n-1} = \frac{ba^{n-1}}{(a+b)^n}, \forall n \in \mathbb{N}_+$$

此分布相当于成功率 $p = \frac{a}{a+b}$ 的几何分布，直接利用第7题的公式，得

$$E(X) = \frac{a+b}{a}$$

6 极端的期望

存在. 反例如下:

样本空间 $\Omega = \{\omega_1, \omega_2\}$, 两样本点概率分别为0.99和0.01.
且有:

$$\begin{aligned}X(\omega_1) &= 1 \\X(\omega_2) &= 10^{100} \\Y(\omega_1) &= Y(\omega_2) = 2\end{aligned}$$

于是:

$$\begin{aligned}E(X) &= 1 \times 0.99 + 10^{100} \times 0.01 = 0.99 + 10^{98} \\E(Y) &= 2\end{aligned}$$

故有

$$E(X) > 100E(Y)$$

然而只要 ω_1 发生, 就有 $Y > X$, 其概率为0.99.

7 几何分布

7.1 分布

第一次试验成功出现在第 m 次相当于前 $m-1$ 次全部失败而第 m 次成功, 其后的情况无所谓. 故

$$P(X = m) = p(1-p)^{m-1}$$

其中 m 为正整数.

X 的分布表:

X	1	2	3	...	m	...
P	p	$p(1-p)$	$p(1-p)^2$...	$p(1-p)^{m-1}$...

7.2 期望

$$\begin{aligned}E(X) &= \sum_{i=1}^{\infty} ip(1-p)^{i-1} \\&= \sum_{i=0}^{\infty} (i+1)p(1-p)^i\end{aligned}$$

于是

$$(1-p)E(X) = \sum_{i=1}^{\infty} ip(1-p)^i$$

两式相减, 有

$$\begin{aligned}pE(X) &= p + \sum_{i=1}^{\infty} p(1-p)^i \\&= p \sum_{i=0}^{\infty} (1-p)^i \\&= p \times \frac{1}{1-(1-p)} \\&= 1\end{aligned}$$

于是

$$E(X) = \frac{1}{p}$$

7.3 方差

由第4题,

$$Var(X) = E(X^2) - E^2(X)$$

先计算第二项,

$$E^2(X) = \frac{1}{p^2}$$

再计算第一项,

$$E(X^2) = \sum_{i=1}^{\infty} i^2 p (1-p)^{i-1}$$

定义函数项级数

$$f(x) = \sum_{i=1}^{\infty} i^2 x^{i-1}$$

则

$$\int_0^x f(t) dt = \sum_{i=1}^{\infty} i x^i$$

$$\frac{\int_0^x f(t) dt}{x} = \sum_{i=1}^{\infty} i x^{i-1}$$

$$\int_0^x \frac{\int_0^s f(t) dt}{s} ds = \sum_{i=1}^{\infty} x^i = \frac{x}{1-x}$$

于是

$$\begin{aligned} f(x) &= \frac{d}{dx} \left(x \frac{d}{dx} \left(\frac{x}{1-x} \right) \right) \\ &= \frac{d}{dx} \left(\frac{x}{(1-x)^2} \right) \\ &= \frac{1+x}{(1-x)^3} \end{aligned}$$

于是

$$E(X^2) = p f(1-p) = p \times \frac{2-p}{p^3} = \frac{2-p}{p^2}$$

故

$$Var(X) = E(X^2) - E^2(X) = \frac{1-p}{p^2}$$

8 市场调查

令 X 为曾通过该平台购买商品的人数, 则

$$X \sim B(25, 0.6)$$

以下数据根据二项分布公式经计算机计算得到.

8.1

$$P(X \geq 15) = 0.5858$$

8.2

$$P(X > 20) = 0.0095$$

8.3

$$P(X < 10) = 0.0132$$

9 二项分布

9.1 期望

$$\begin{aligned}
 E(X) &= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\
 &= 0 + \sum_{k=1}^n \frac{k n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
 &= n \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\
 &= np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k} \\
 &= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-1-k)!} p^k (1-p)^{n-1-k} \\
 &= np(p + (1-p))^{n-1} \\
 &= np
 \end{aligned}$$

9.2 方差

$$Var(X) = E(X^2) - E^2(X)$$

后一项有

$$E^2(X) = n^2 p^2$$

考虑前一项

$$\begin{aligned}
 E(X^2) &= \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} \\
 &= \sum_{k=1}^n \frac{k^2 n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
 &= np \sum_{k=1}^n \frac{k(n-1)!}{(k-1)!(n-k)!} p^{k-1} (1-p)^{n-k} \\
 &= np \sum_{k=0}^{n-1} \frac{(k+1)(n-1)!}{k!(n-1-k)!} p^k (1-p)^{n-1-k} \\
 &= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-1-k)!} p^k (1-p)^{n-1-k} + np \sum_{k=1}^{n-1} \frac{k(n-1)!}{k!(n-1-k)!} p^k (1-p)^{n-1-k} \\
 &= np + n(n-1)p^2 \sum_{k=1}^{n-1} \frac{(n-2)!}{(k-1)!(n-1-k)!} p^{k-1} (1-p)^{n-1-k} \\
 &= np + n(n-1)p^2 \sum_{k=0}^{n-2} \frac{(n-2)!}{k!(n-2-k)!} p^k (1-p)^{n-2-k} \\
 &= np + n(n-1)p^2(p + (1-p))^{n-2} \\
 &= np + n^2 p^2 - np^2
 \end{aligned}$$

故

$$Var(X) = np - np^2 = np(1-p)$$

10 标记重捕法

10.1

总共有 $\binom{N}{n}$ 种捞鱼的情形，其中捞到 m 条有记号（即 M 中选 m 同时 $N-M$ 中选 $n-m$ ）的可能共有 $\binom{M}{m}\binom{N-M}{n-m}$ 种，故

$$P(X = m) = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}$$

10.2

假定放回后充分混合，且标记并不影响捕捉难易程度，则被捕捞上来的鱼在所有鱼和所有标记鱼两个总体中占比应基本相同（重捕的比例独立于是否有标记），即：

$$\frac{n}{N} = \frac{m}{M}$$

故

$$N = \frac{Mn}{m}$$

10.3

共有 N 条鱼时的概率为

$$P(X = m)|_N = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}} = \frac{M!(N-M)!n!(N-n)!}{N!m!(M-m)!(n-m)!(N-M-n+m)!}$$

故

$$\frac{P(X = m)|_{N+1}}{P(X = m)|_N} = \frac{(N-M+1)(N-n+1)}{(N+1)(N-M-n+m+1)}$$

此比值总是正数，下面令此比值大于1，有

$$(N-M+1)(N-n+1) > (N+1)(N-M-n+m+1)$$

$$N^2 - (M+n-2)N + (Mn-M-n+1) > N^2 - (M+n-m-2)N + (-M-n+m+1)$$

$$mN < Mn - m$$

$$N < \frac{Mn}{m} - 1$$

即，当 $N < \frac{Mn}{m} - 1$ 时， $N+1$ 比 N 概率增加，而 $N > \frac{Mn}{m} - 1$ 时， $N+1$ 比 N 概率减小，而 $N = \frac{Mn}{m} - 1$ 和 $N = \frac{Mn}{m}$ 的概率相等。

然而考虑到 $\frac{Mn}{m}$ 并不总是整数，不是整数时上述概率表达式没有意义，使得上述概率最大的 N 应当取区间 $[\frac{Mn}{m} - 1, \frac{Mn}{m}]$ 中的整数，即 $\lfloor \frac{Mn}{m} \rfloor$ 。

此结果几乎与第二问的估计值相同。

10.4

极大似然估计是在某随机变量的值可以观测确定，但其一个影响因素不可知的情况下，对该影响因素的一种估计策略：认为使得此随机变量取到给定的观测值的概率最高的值就是该影响因素的估计值。

我的理解是，在对待估计值的大致分布情况一无所知的情况下，即认为，（至少是概率极大的值附近的）各种可能值都等可能出现，极大似然估计是有道理的。然而，对某结果的成因的最常见估计策略是贝叶斯公式，假如我们对该待测值的可能分布有一个粗略的估计，例如在某个值附近形成近似的正态分布，即：我们已知各种可能取值的先验概率，那么这种情况下条件概率最大的条件未见得就是最可能的，而应是先验概率与条件概率之积最大（即后验概率最大）的那个值。

11 计算机实验：二项分布

11.1

X取15时有最大概率.

11.2

期望值为 $\mu = 25 \times 0.6 = 15$ ，即最大概率的X值.

11.3

方差为 $\sigma^2 = 25 \times 0.6 \times 0.4 = 6$.

11.4

$P(\mu - 2\sigma < x < \mu + 2\sigma) = P(3 < X < 27) = 1 - 9.54 \times 10^{-7}$.

11.5 所用代码

```
> x<-dbinom(0:25,25,0.6)
> x
[1] 1.125900e-10 4.222125e-09 7.599824e-08 8.739798e-07 7.210333e-06 4.542510e-05 2.271255e-04 9.247253e-04
[9] 3.120948e-03 8.842685e-03 2.122244e-02 4.340955e-02 7.596671e-02 1.139501e-01 1.465072e-01 1.611579e-01
[17] 1.510856e-01 1.199797e-01 7.998648e-02 4.420305e-02 1.989137e-02 7.104062e-03 1.937471e-03 3.790705e-04
[25] 4.738381e-05 2.843029e-06
> plot(0:25,x,xlab="x",ylab="P",main="B(25,0.6)")
```

11.6 分布图

