Material didático preparado pelo professor Maurício Mota para a disciplina CC0282- Probabilidade I ministrada no semestre 2021.1.

1 Distribuição Uniforme

1.1 Introdução

1.2 Definição

Uma variável aleatória discreta X é dita possuir *distribuição uniforme* de parâmetro $N \in \mathbb{N}^*$ se sua função de probabilidade ($f.p.$) é da forma:

$$f(x) = \frac{1}{N} I_{\{1,2,\ldots,N\}}(x) = \begin{cases} \frac{1}{N}, & \text{se } x = 1, 2 \ldots, N \\ 0, & \text{para outros valores de } x \end{cases} \tag{1}$$

Notação: $X \sim Unif(N)$.

Observação: Lê-se a notação acima do seguinte modo: X segue distribuição $Uniforme$ de parâmetro N.

Os gráficos da função de probabilidade de $X \sim Unif(N)$, para alguns valores do parâmetro N, estão na Figura 1.

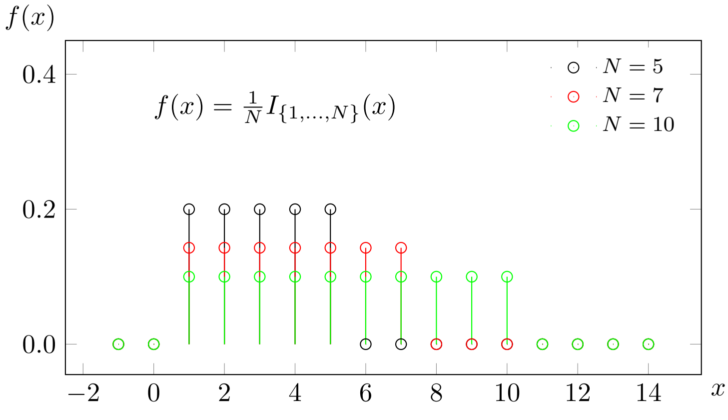

Figura 1: Gráfico da Função de Probabilidade Uniforme

1.3 Propriedades da função de probabilidade

Fato 1. A expressão (1) é realmente uma legítima função de probabilidade.

Prova: Deve-se verificar que

i. $f(x) > 0,\ x \in A$.

ii. $\displaystyle\sum_A f(x) = 1$,

sendo $A = \{x \in R \mid f(x) > 0\}$ o suporte da distribuição de X. Como $A = \{1, 2, \ldots, N\}$ é o suporte e $f(x) = \dfrac{1}{N} > 0, x \in A$. Assim $f(x) \geq 0$. A segunda propriedade nos diz que a soma dos valores das probabilidades para os pontos do suporte é 1. Assim ,

$$\sum_A f(x) = \sum_{x=1}^{N} \frac{1}{N} = N\,\frac{1}{N} = 1 \qquad \blacksquare$$

Fato 2. A distribuição uniforme é simétrica em torno do ponto $c = \frac{N+1}{2}$.

Prova:

Deve-se verificar que $f(c-x) = f(c+x),\ \forall x$. Vai-se analisar inicialmente o caso N ímpar. Considere o conjunto $B = \left\{1, 2, \ldots, \frac{N-1}{2}\right\}$ no qual qualquer elemento tem probabilidade não nula. Seja $x \in B$, assim $c+x$ e $c-x$ pertencem ao suporte da distribuição e portanto $f(c-x) = f(c+x) = \dfrac{1}{N}, \forall x \in B$. Se $x \in B^c$ então $c+x$ e $c-x$ não pertencem ao suporte da distribuição e portanto $f(c-x) = f(c+x) = 0,\ \forall x \in B^c$. Assim para todo x real $f(c-x) = f(c+x)$, e a simetria está provada.

Para o caso N par, considere o conjunto $B = \left\{\frac{1}{2}, \frac{3}{2}, \ldots, \frac{N-1}{2}\right\}$ e portanto $c+x$ e $c-x$ pertencem ao suporte para qualquer elemento x de B e portanto tem a mesma probabilidade. Para x fora de B as probabilidades são nulas. Isso completa a prova $\blacksquare$

1.4 Momentos em relação à origem

Vai-se calcular os quatro primeiros momentos em relação à origem. Os seguintes somatórios serão úteis:

$\mathcal{S}1.$ $\displaystyle\sum_{x=1}^{N} x = \frac{N(N+1)}{2}$;

$\mathcal{S}2.$ $\displaystyle\sum_{x=1}^{N} x^2 = \frac{N(N+1)(2N+1)}{6}$;

$\mathcal{S}3.$ $\displaystyle\sum_{x=1}^{N} x^3 = \left[\frac{N(N+1)}{2}\right]^2$; e

$\mathcal{S}4.$ $\displaystyle\sum_{x=1}^{N} x^4 = \frac{N(N+1)(2N+1)(3N^2+3N-1)}{30}$.

1.4.1 Primeiro Momento

Fato 3. Se $X \sim Unif(N)$, então $E(X) = \dfrac{N+1}{2}$.

<u>*Prova*</u>: Lembrando da propriedade $\mathcal{S}1$, então

$$E(X) = \sum_{x=1}^{N} x\frac{1}{N} = \frac{1}{N}\sum_{x=1}^{N} x = \frac{1}{N}\frac{N(N+1)}{2} = \frac{N+1}{2} \qquad \blacksquare$$

1.4.2 Segundo Momento

Fato 4. Se $X \sim Unif(N)$, então $E(X^2) = \dfrac{(N+1)(2N+1)}{6}$.

<u>*Prova*</u>:

$$E(X^2) = \sum_{x=1}^{N} x^2\frac{1}{N} = \frac{1}{N}\sum_{x=1}^{N} x^2 = \frac{1}{N}\frac{N(N+1)(2N+1)}{6} = \frac{(N+1)(2N+1)}{6} \qquad \blacksquare$$

1.4.3 Terceiro Momento

Fato 4. Se $X \sim Unif(N)$, então $E(X^3) = \dfrac{N(N+1)^2}{4}$.

<u>*Prova*</u>:

$$E(X^3) = \sum_{x=1}^{N} x^3\frac{1}{N} = \frac{1}{N}\sum_{x=1}^{N} x^3 = \frac{1}{N}\left[\frac{N(N+1)}{2}\right]^2 = \frac{N(N+1)^2}{4} \qquad \blacksquare$$

1.4.4 Quarto Momento

Fato 5. Se $X \sim Unif(N)$, então $E(X^4) = \dfrac{(N+1)(2N+1)(3N^2+3N-1)}{30}$.

<u>*Prova*</u>:

$$E(X^4) = \sum_{x=1}^{N} x^4\frac{1}{N} = \frac{1}{N}\sum_{x=1}^{N} x^4 = \frac{1}{N}\frac{N(N+1)(2N+1)(3N^2+3N-1)}{30},$$

assim,

$$E(X^4) = \frac{(N+1)(2N+1)(3N^2+3N-1)}{30} \qquad \blacksquare$$

1.5 Momentos Centrais

Vai-se agora calcular a variância (segundo momento central), o terceiro e o quarto momentos centrais de $X \sim Unif(N)$.

1.5.1 Segundo Momento Central

Fato 6. Se $X \sim Unif(N)$, então $\mu_2 = E\left[(X-\mu)^2\right] = \sigma^2 = \dfrac{N^2-1}{12}$.

<u>*Prova*</u>:

$$E\left[(X-\mu)^2\right] = E(X^2) - E^2(X),$$

assim,

$$\mu_2 = \frac{(N+1)(2N+1)}{6} - \left[\frac{N+1}{2}\right]^2 = \frac{(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4},$$

que pode ser escrito como

$$\mu_2 = \frac{2(N+1)(2N+1)}{12} - \frac{3(N+1)^2}{12} = \frac{(N+1)(4N+2-3N-3)}{12},$$

colocando $(N+1)$ em evidência tem-se

$$\mu_2 = \frac{(N+1)(4N+2-3N-3)}{12} = \frac{(N+1)(N-1)}{12} = \frac{N^2-1}{12} \qquad \blacksquare$$

1.5.2 Terceiro Momento Central

Fato 7. Se $X \sim Unif(N)$, então $\mu_3 = E\left[(X-\mu)^3\right] = 0$.

<u>*Prova*</u>:

$$E\left[(X-\mu)^3\right] = E(X^3) - 3E(X)E(X^2) + 2E^3(X) = E(X^3) - 3\mu E(X^2) + 2\mu^3,$$

assim,

$$\mu_3 = \frac{N(N+1)^2}{4} - 3\frac{(N+1)}{2}\,\frac{(N+1)(2N+1)}{6} + 2\frac{(N+1)^3}{8} =,$$

que pode ser escrito como

$$\mu_3 = \frac{N(N+1)^2}{4} - \frac{(N+1)^2(2N+1)}{4} + \frac{(N+1)^3}{4} =,$$

colocando $(N+1)^2$ em evidência tem-se

$$\mu_3 = \frac{(N+1)^2(N-2N-1+N+1)}{4} = 0,$$

resultado esse esperado pois a distribuição uniforme discreta é simétrica em torno de sua média. Assim qualquer momento central de ordem ímpar é nulo $\blacksquare$

1.5.3 Quarto Momento Central

Fato 8. Se $X \sim Unif(N)$, então $\mu_4 = E\left[(X-\mu)^4\right] = \dfrac{(N^2-1)(3N^2-7)}{240}$.

Prova:

$$\mu_4 = E(X^4) - 4E(X^3)E(X) + 6E(X^2)E^2(X) - 3E^4(X) =,$$

assim,

$$\begin{aligned}
\mu_4 &= \frac{(N+1)(2N+1)(3N^2+3N-1)}{30} - 4\frac{N(N+1)^2}{4}\,\frac{N+1}{2} + 6\frac{(N+1)(2N+1)}{6}\,\frac{(N+1)^2}{4} - 3\frac{(N+1)^4}{16}\\
&= \frac{(N+1)(2N+1)(3N^2+3N-1)}{30} - \frac{N(N+1)^3}{2} + \frac{(2N+1)(N+1)^3}{4} - 3\frac{(N+1)^4}{16}\\
&= (N+1)\left[\frac{(2N+1)(3N^2+3N-1)}{30} - \frac{N(N+1)^2}{2} + \frac{(2N+1)(N+1)^2}{4} - \frac{3(N+1)^3}{16}\right]\\
&= (N+1)\left[\frac{(2N+1)(3N^2+3N-1)}{30} - (N+1)^2\left[\frac{N}{2} - \frac{(2N+1)}{4} + \frac{3(N+1)}{16}\right]\right]\\
&= (N+1)\left[\frac{(2N+1)(3N^2+3N-1)}{30} - (N+1)^2\left[\frac{8N}{16} - \frac{(8N+4)}{16} + \frac{3(N+1)}{16}\right]\right]\\
&= (N+1)\left[\frac{(2N+1)(3N^2+3N-1)}{30} - (N+1)^2\left[\frac{8N-8N-4+3N+3}{16}\right]\right]\\
&= (N+1)\left[\frac{(2N+1)(3N^2+3N-1)}{30} - \frac{(N+1)^2(3N-1)}{16}\right]\\
&= (N+1)\left[\frac{8(2N+1)(3N^2+3N-1) - 15(N+1)^2(3N-1)}{240}\right]\\
&= (N+1)\left[\frac{8(6N^3+9N^2+N-1) - 15(3N^3+15N^2+N-1)}{240}\right]\\
&= (N+1)\left[\frac{3N^3-3N^2-7N+7}{240}\right] = (N+1)\left[\frac{3N^2(N-1)-7(N-1)}{240}\right]\\
&= (N+1)(N-1)\left(\frac{3N^2-7}{240}\right) = \frac{(N^2-1)(3N^2-7)}{240} \qquad \blacksquare
\end{aligned}$$

1.6 Coeficiente de Assimetria

Fato 9. O coeficiente de assimetria de $X \sim Unif(N)$ é $\alpha_3 = 0$.

Prova: Como o terceiro momento centra de uma variável aleatória uniforme é zero, então

$$\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}} = 0,$$

e assim nada se pode afirmar sobre a simetria da distribuição Uniforme usando o coeficiente de assimetria. Mas pela definição provou-se que ela é simétrica em torno de $\frac{N+1}{2}$.

1.7 Coeficiente de Curtose

Fato 10. O coeficiente de curtose de $X \sim Unif(N)$ é $\alpha_4 = \dfrac{3}{5}\dfrac{(3N^2-7)}{N^2-1}$.

Prova:

$$\alpha_4 = \frac{\mu_4}{\mu_2^2} = \frac{\dfrac{(N^2-1)(3N^2-7)}{240}}{\dfrac{(N^2-1)^2}{144}} = \frac{3}{5}\frac{(3N^2-7)}{N^2-1}.$$

Como

$$\alpha_4 = \frac{3}{5}\frac{(3N^2-7)}{N^2-1} = \frac{3}{5}\frac{(3N^2-3-4)}{N^2-1} = \frac{3}{5}\frac{[3(N^2-1)-4]}{N^2-1},$$

assim,

$$\alpha_4 = \frac{3}{5}\left(3 - \frac{4}{N^2-1}\right) < 3.$$

Portanto pode-se classificar a distribuição de Uniforme discreta quanto à curtose como platicúrtica.

1.8 Momentos Fatoriais

Fato 11. Se $X \sim Unif(N)$, então

$$E\left(X_{[r]}\right) = \begin{cases} 0, & \text{se } r \geq N, \\ \dfrac{(N+1)(N-1)!}{(r+1)(N-r)!}, & \text{se } r \leq N-1. \end{cases}$$

Prova:

$$\mu_{[r]} = E\left[X(X-1)\dots(X-r+1)\right] = \sum_{x=1}^{N} x(x-1)\dots(x-r+1)\frac{1}{N}.$$

Se $r \geq N$ todas as parcelas se anulam logo $\mu_{[r]} = 0$. Para $1 \leq r \leq (N-1)$ tem-se assim,

$$\begin{aligned}
\mu_{[r]} &= \frac{1}{N}\sum_{x=r}^{N} x(x-1)\dots(x-r+1) \\
&= \frac{1}{N}\sum_{x=r}^{N}\frac{x!}{(x-r)!} \\
&= \frac{1}{N}\sum_{x=r}^{N}\frac{r!x!}{r!(x-r)!} \\
&= \frac{r!}{N}\sum_{x=r}^{N}\binom{x}{r} \\
&= \frac{r!}{N}\binom{N+1}{r+1} \\
&= \frac{(N+1)!}{N(r+1)(N-r)!} = \frac{(N+1)(N-1)!}{(r+1)(N-r)!} \qquad \blacksquare
\end{aligned}$$

1.9 Função Geradora de Probabilidades

Fato 12. Se $X \sim Unif(N)$, então

$$G(t) = \begin{cases} 1, & \text{se } \ t = 1 \\ \dfrac{t(1-t^N)}{N(1-t)}, & \text{se } \ t \neq 1. \end{cases}$$

Prova: Para $t = 1$ tem-se $G(t) = E(1^X) = E(1) = 1$. Para $t \neq 1$ tem-se:

$$G(t) = E(t^X) = \sum_{x=1}^{N} t^x \frac{1}{N} = \frac{1}{N} \sum_{x=1}^{N} t^x = \frac{1}{N} \frac{t - t^{N+1}}{(1-t)} = \frac{t(1-t^N)}{N(1-t)} \qquad \blacksquare$$

1.10 Função Geradora de Momentos

Fato 13. Se $X \sim Unif(N)$, então

$$M(t) = \begin{cases} 1, & \text{se } \ t = 0 \\ \dfrac{e^t(1-e^{Nt})}{N(1-e^t)}, & \text{se } \ t \neq 0. \end{cases}$$

Prova: Para $t = 0$ tem-se $M(0) = E(e^{0^X}) = E(1) = 1$. Para $t \neq 0$ tem-se:

$$M(t) = E(e^{tX}) = \varphi(e^t),$$

se $e^t \neq 1$, isto é, se $t \neq 0$ tem-se,

$$M(t) = \frac{e^t(1-e^{Nt})}{N(1-e^t)} \qquad \blacksquare$$

1.11 Função de Distribuição

Fato 14. A função de distribuição de $X \sim Unif(N)$

$$F(x) = \frac{[x]}{N} I_{[\ 1,\ N)}(x) + I_{[\ N,\ \infty)}(x),$$

onde $[x]$ é o maior inteiro que não ultrapassa a x.

Prova:

Para $x < 1$ tem-se $F(x) = 0$, para $x \geq N$ tem-se $F(x) = 1$ e para $1 \leq x < N$, tem-se

$$F(x) = P(X \leq x) = P(X \leq [x]) = \sum_{x=1}^{[x]} \frac{1}{N} = \frac{[x]}{N} \qquad \blacksquare$$

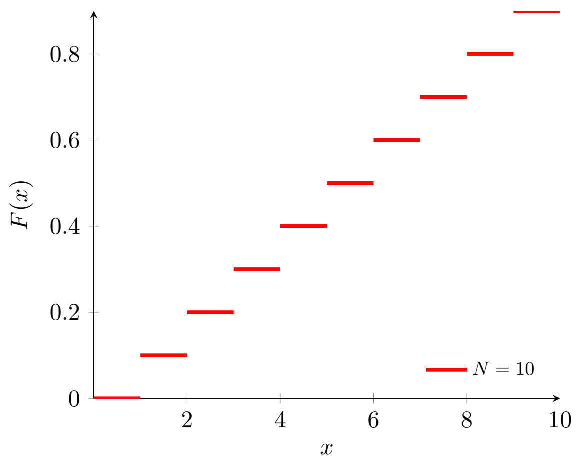

Figura 2: Gráfico da Função de Distribuição Uniforme

O gráfico da função de distribuição de $X \sim Unif(10)$ está na Figura 2.

1.12 Função de Sobrevivência

A função de sobrevivência de $X \sim Unif(N)$

$$S(x) = I_{(-\infty,\,1)}(x) + \frac{n - [x]}{N} I_{[\,1,\,N)}(x).$$

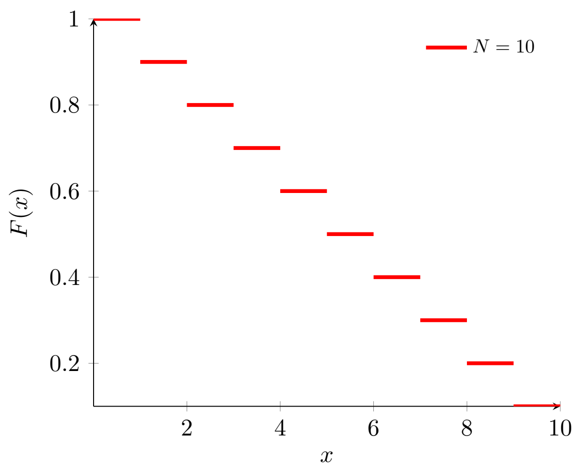

Figura 3: Gráfico da Função de Sobrevivência Uniforme

Os gráfico da função de sobrevivência de $X \sim Unif(10)$ está na Figura 3.

1.13 Mediana

Fato 15. A mediana da distribuição de $X \sim Unif(N)$

$$med = \begin{cases} \dfrac{N+1}{2}, & \text{se } N \text{ é ímpar.} \\ \\ \dfrac{N}{2}, & \text{se } N \ \text{ é par.} \end{cases}$$

Prova:

Usando o fato de que a mediana é o menor valor x da variável que satisfaz $F(x) \geq 1/2$ o resultado é imediato, pois para o caso N ímpar,

$$F\left(\frac{N-1}{2}\right) = \frac{1}{2} - \frac{1}{2N} < \frac{1}{2},$$

e

$$F\left(\frac{N+1}{2}\right)) = \frac{1}{2} + \frac{1}{2N} > \frac{1}{2},$$

portanto, $med = \dfrac{N+1}{2}$. Por outro lado, para o caso N par,

$$F\left(\frac{N}{2} - 1\right) = \frac{1}{2} - \frac{1}{2N} < \frac{1}{2},$$

e

$$F\left(\frac{N}{2}\right) = \frac{1}{2},$$

portanto $med = \dfrac{N}{2}$ $\blacksquare$.

1.14 Moda

Fato 16. A distribuição de $X \sim Unif(N)$ é amodal pois todos os valores do suporte são equiprováveis.

1.15 Coeficiente de Variação

Fato 17. O coeficiente de variação de $X \sim Unif(N)$

$$CV = \frac{\sigma}{\mu} = \sqrt{\frac{N-1}{3(N+1)}}.$$

Prova:

$$CV = \frac{\sigma}{\mu} = \frac{2}{N+1}\sqrt{\frac{N^2-1}{12}} = \sqrt{\frac{4}{(N+1)^2}\frac{N^2-1}{12}} = \sqrt{\frac{N-1}{3(N+1)}} \quad \blacksquare$$

1.16 Transformação

Sejam $X \sim Unif(N)$ e $Y \sim Unif(N)$ independentes.

Mostre que:

i
$$P(Y = X) = \frac{1}{N}.$$

Prova:

$$\begin{aligned} P(Y = X) &= \sum_{x=1}^{N} P(X = x, Y = x) \\ &= \sum_{x=1}^{N} P(X = x)P(Y = x) \\ &= \sum_{x=1}^{N} \frac{1}{N^2} \\ &= \frac{N}{N^2} \\ &= \frac{1}{N}. \end{aligned}$$

ii
$$P(Y > X) = \frac{N-1}{2N}.$$

$$
\begin{aligned}
P(Y > X) &= \sum_{x=1}^{N} P(X = x, Y > x) \\
&= \sum_{x=1}^{N} P(X = x)P(Y > x) \\
&= \sum_{x=1}^{N} \sum_{y=x+1}^{N} \frac{1}{N} \\
&= \sum_{x=1}^{N} \frac{N-x}{N} \\
&= \frac{1}{N^2} \sum_{x=1}^{N} (N-x) \\
&= \frac{1}{N^2} \left[N^2 - \sum_{x=1}^{N} x \right] \\
&= \frac{1}{N^2} \left[N^2 - \frac{N(N+1)}{2} \right] \\
&= \frac{1}{N^2} \left[N^2 - \frac{N(N-1)}{2} \right] \\
&= \frac{N-1}{2N}.
\end{aligned}
$$

Qual a função de probabilidade de?

a. $S = X + Y$,

Solução: O suporte de S é dado por:

$$A = \{2, 3, \ldots, 2N-1, 2N\}.$$

Assim,

Para $s \in A$ temos:

$$
\begin{aligned}
P(S = s) &= \sum_{x} P(X = x, Y = s - x) \\
&= \sum_{L_i}^{L_s} P(X = x)P(Y = s - x) \\
&= \sum_{L_i}^{L_s} \frac{1}{N^2} \\
&= \frac{L_s - L_i + 1}{N^2}.
\end{aligned}
$$

Vamos calcular os limites do somatório:

temos que:

$$1 \leq x \leq N \tag{2}$$

e

$$1 \leq s - x \leq N$$

,

Vamos trabalhar com esta última desigualdade temos:

$$-N \leq x - s \leq -1$$

e

$$s - N \leq x \leq s - 1 \tag{3}$$

assim

$$L_i = max(s - N, 1) \leq x \leq min(s - 1, N) = L_s. \tag{4}$$

A função de probabilidade de $S = X + Y$ é dada por

$$P(S = s) = \frac{L_s - L_i + 1}{N^2} = \frac{min(s - 1, N) - max(s - N, 1) + 1}{N^2}; I_{\{2,\ldots,2N\}}(s).$$

Uma maneira alternativa

$$P(S = s) = \frac{s - 1}{N^2}; I_{\{2,\ldots,N\}}(s) + \frac{2N - s + 1}{N^2}; I_{\{N+1,\ldots,2N\}}(s).$$

b.

$$D = X - Y.$$

c.

$$U = min(X, Y).$$

d.

$$V = max(X, Y).$$

1.17 Simulação

Vai-se discutir agora a simulação de uma amostra aleatória de tamanho n de $X \sim Unif(N)$ usando o pacote estatístico R. Como exemplo vai-se simular uma amostra de tamanho 25 de $X \sim Unif(N = 5)$. Isto é feito através dos comandos:

```
> ### Seja A={1,2,3,4,5}, o suporte da distribuição
> A=1:5
>  dados=sample(A,25,replace=T)
> table(dados)
dados 1 2 3 4 5
      6 5 3 3 8
#Assim o valor 1 ocorreu 6 vezes, o 2 ocorreu 5 vezes e assim por diante.
```

Agora se vai simular 1000 lançamentos de uma moeda honesta. A saída do R é dada abaixo.

```
> ##Simular um lançamento de uma moeda honesta 1000 vezes,1=cara,0=coroa
> ### Seja A={0,1}
> A=0:1
> dados=sample(A,1000,replace=T);
> table(dados)
dados
  0   1
525 475
```

Assim dos 1000 lançamentos da moeda aconteceram 475 caras e 525 coroas. Agora se vai simular 1200 lançamentos de um dado honesto. A saída do R é dada abaixo.

```
> ##Simular um lancamento de um dado honesto 2000 vezes
> ### Seja A={1,2,3,4,5,6}
> A=1:6
>  dados=sample(A,1200,replace=T)
> table(dados)
dados
  1   2   3   4   5   6
185 226 208 170 216 195
```

As frequências estão bem próximas da frequência esperada 200. Também pode-se usar o `R` para a amostragem sem reposição. De um grupo de 20 alunos numerados de 1 a 20, 5 formarão uma comissão. Vamos escolher com a ajuda do `R`!!!

```
> U=1:20  #universo dos alunos
> comissao=sample(U,5)  #sem reposição
> comissao
[1] 11 17 18 15  9
```

A comissão será formada pelos alunos de numeração: 9, 11,15,17 e 18.

1.18 Transformações para a uniforme

Caso 1. Seja $A = \{a, a+1, \ldots, b-1\}$ o suporte de uma variável aleatória discreta onde todos os $(b-1-a+1) = b-a$ pontos são equiprováveis. A f.p. de Y é dada por:

$$f(y) = \frac{1}{b-a} I_{\{a\ ,\ a+1\ ,\ldots\ ,\ b-1\}}(y).$$

Considere a transformação $X = Y - a + 1$. O suporte de X é o conjunto $\{1, 2, \ldots, b-a\}$. Assim a f.p. de X é dada por

$$f_X(x) = P(X = x) = P(Y - a + 1 = x = P(Y = x + a - 1),$$

que fica assim após a substituição

$$f_X(x) = \frac{1}{b-a} I_{\{a\ ,\ a+1\ ,\ldots\ ,\ b-1\}}(x + a - 1) = \frac{1}{b-a} I_{\{1\ ,\ 2\ ,\ldots\ ,\ b-a\}}(x).$$

Assim $X = Y - a + 1 \sim Unif(N = b - a)$. Este fato é bastante útil pois a média e a variância de Y são facilmente calculadas a partir da média e da variância de X.

Se $X \sim Unif(b-a)$, então

$$E(X) = \frac{b-a+1}{2}$$

e a média de Y é dada por:

$$E(Y) = E(X + a - 1) = E(X) + a - 1 = \frac{b-a+1}{2} + a - 1 = \frac{a+b-1}{2}, \text{ e}$$

$$Var(Y) = V(X + a - 1) = Var(X) = \frac{(b-a)^2 - 1}{12}.$$

Quaisquer outros momentos de Y podem se calculados a partir dos momentos de X.

Caso 2. Considere uma variável aleatória Y com suporte $A = \{y_1, y_2, \ldots, y_N\}$, em que $y_j = a + jh,\ j = 1, 2, \ldots, N$. Assim a transformação $X = \dfrac{Y-a}{h}$ tem distribuição uniforme sobre $A = \{1, 2, \ldots, N\}$. Logo,

$$E(Y) = E(a + hX) = hE(X) + a = h\frac{N+1}{2} + a = \frac{h(N+1)}{2}, \text{ e}$$

$$Var(Y) = V(hX + a) = h^2 Var(X) = \frac{h^2(N^2-1)}{12}.$$

Caso 3. Considere uma variável aleatória Y com suporte equiprovável $A = \{y_1, y_2, \ldots, y_N\}$, sem nenhuma lei de formação entre os elementos do suporte. A função de probabilidade de Y é dada por:

$$f(y) = \frac{1}{N} I_A(y).$$

A esperança de Y é dada por:

$$E(Y) = \frac{\sum_{i=1}^{N} y_i}{N} = \bar{y}.$$

A $E(Y^2)$ é dada por:

$$E(Y^2) = \frac{\sum_{i=1}^{N} y_i^2}{N}.$$

A variância de Y é dada por:

$$var(Y) = E((Y - \bar{y})^2) = \frac{\sum_{i=1}^{N} (y_i - \bar{y})^2}{N} = \frac{\sum_{i=1}^{N} y_i^2 - N\bar{y}^2}{N}.$$

1.19 Entropia

A entropia é uma das características mais úteis das distribuições. Ela foi introduzida por Shanon e é definida como:

$$H(X) = E(-lnX) = -\sum_{A} \ln x\, P(X = x).$$

Fato 18. A entropia da distribuição uniforme $X \sim Unif(N)$ é dada por:

$$H(X) = E(-\ln X) = -\sum_{x=1}^{N} \ln x \frac{1}{N} = -\frac{\ln N!}{N}.$$

1.20 Estimação do parâmetro N

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de $X \sim Unif(N)$. Os seguintes estimadores podem ser usados para se estimar N:

a. Estimador 1: $\hat{N}_1 = Y_n = Max(X_1, X_2, \ldots, X_n)$;

b. Estimador 2: $\hat{N}_2 = \dfrac{n+1}{n} Y_n = \dfrac{n+1}{n} Max(X_1, X_2, \ldots, X_n)$; e

c. Estimador 3: $\hat{N}_3 = 2\bar{X} - 1$, $\bar{X}$ é a média amostral.

$$\bar{X} = \frac{S}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

1.21 Teste de Aderência para a Distribuição Uniforme

Para se verificar se uma determinada população segue uma distribuição uniforme de parâmetro N uma amostra aleatória de tamanho n é retirada e o número de vezes que cada valor do suporte ocorre é anotado gerando uma tabela do tipo:

Ocorrência	1	2	...	N	Total
Frequência observada (O)	O_1	O_2	...	O_N	n

Se a distribuição é uniforme a frequência esperada(E) de cada casela é dada por $E = \dfrac{n}{N}$. Assim completa-se a tabela:

Ocorrência	1	2	...	N	Total
Frequência observada (O)	O_1	O_2	...	O_N	n
Frequência esperada (E)	E_1	E_2	...	E_N	n

Deve-se comparar as frequências observadas com as esperadas através da estatística:

$$Q = \sum_{i=1}^{N} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2(N-1),$$

e agora, considere a hipótese H_0 : a distribuição é Uniforme de parâmetro N. Para testá-la vai-se usar um nível de significância α e a seguinte regra de decisão: rejeitar H_0 se $Qcal > Qtab(1-\alpha, N-1)$ e nesse caso de uniformidade a estatística do teste é simplificada para :

$$Q = \left(\frac{N}{n} \sum_{i=1}^{N} O_i^2 \right) - n.$$

1.22 Exercícios Resolvidos

1. Suponha que um estudante é selecionado aleatoriamente de um grupo de 6 estudantes para ser líder da turma de Probabilidade I de 2009.1. Suponha que para efeito de sorteio eles sejam numerados

de 1 a 6. Seja X a numeração do estudante escolhido. Pede-se:

a. Identificar a distribuição de X;

Solução: $A = \{1, 2, 3, 4, 5, 6\}$ é o suporte da distribuição e pelo enunciado qualquer ponto do suporte é equiprovável. Portanto X tem distribuição uniforme discreta de parâmetro $N = 6$.

b. Qual a função de probabilidade de X?;

Solução: Como $X \sim Unif(6)$ tem-se $f(x) = \frac{1}{6} I_{\{1,2,3,4,5,6\}}(x)$.

c. Qual a média e a variância de X?;

Solução: Como $X \sim Unif(6)$ tem-se

$$E(X) = \frac{N+1}{2} \text{e} \quad Var(X) = \frac{N^2-1}{12},$$

portanto

$$E(X) = \frac{N+1}{2} = \frac{6+1}{2} = 3,5 \quad \text{e} \quad Var(X) = \frac{6^2-1}{12} = \frac{35}{12}.$$

2. No começo da semana o departamento de esportes de uma grande loja se abastece com um certo número de unidades de um determinado artigo o qual tem uma demanda uniforme de parâmetro 10, isto é,

$$f(x) = \frac{1}{10} I_{\{1,2,3,4,5,6,7,8,9,10\}}(x),$$

onde x é um valor da quantidade procurada X. A empresa tem um lucro líquido de 6 u.m. para cada unidade vendida e uma perda de 2 u.m. para cada uma que não consegue vender na semana. Achar o valor de k o número de unidades que se deve adquirir no início da semana) para maximizar o lucro (L)esperado sabendo que isto acontece apenas uma vez por semana.

Solução: O lucro L é dado por:

$$L = \begin{cases} 6X - 2(k - X) = 8X - 2k & \text{se } X \leq k \\ 6k & \text{se } \quad X > k \end{cases}$$

O valor esperado de $L = h(X)$, $E(L)$ é dado por:

$$E(L) = \sum_{x=1}^{10} h(x) \frac{1}{10} = \frac{1}{10} \left[\sum_{x=1}^{k} h(x) + \sum_{x=k+1}^{10} h(x) \right],$$

assim,

$$E(L) = \frac{1}{10} \left[\sum_{x=1}^{k} (8x - 2k) + \sum_{x=k+1}^{10} 6k \right] = \frac{1}{10} \left[8 \sum_{x=1}^{k} x - 2k^2 + 6k \sum_{x=k+1}^{10} 1 \right],$$

logo,

$$E(L) = \frac{1}{10} \left[8 \frac{k(k+1)}{2} - 2k^2 + 6k(10-k) \right] = \frac{2(16k - k^2)}{5}.$$

Ocorrência	1	2	3	4	5	6
Frequência	188	216	209	191	181	215

O lucro líquido é uma função quadrática de k com $a = -2/5,\ b = 32/5,\ c = 0$. Como a negativo, o ponto que maximiza o valor esperado do lucro é $k = \frac{-b}{2a} = \frac{-32/5}{-4/5} = 8$. Este valor será encontrado agora usando o R. Veja como se faz.

```
> k=1:10
> L=2*(16*k-k^2)/5
> rbind(k,L)
  [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
k    1  2.0  3.0  4.0    5    6  7.0  8.0  9.0    10 L    6 11.2
15.6 19.2   22   24 25.2 25.6 25.2    24
> ###para achar direto
> max(L)
[1] 25.6
> which(L==max(L))
[1] 8
```

3. Um dado foi lançado 1200 vezes, com os seguintes resultados:

   Teste a hipótese que o dado é balanceado. Isto é teste a hipótese de que X , a face do dado que aparece em uma jogada, é uniforme.

   Solução: Se o dado é balanceado, a probabilidade de aparecer cada face é $1/6$ e esperamos que cada face apareça 200 vezes. Vai-se acrescentar esse dado na tabela:

Ocorrência	1	2	3	4	5	6
Frequência Observada	188	216	209	191	181	215
Frequência Esperada	200	200	200	200	200	200

Seja

$$Q = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2(N-1).$$

Considere a hipótese H_0 : a distribuição é Unif(6). Para testá-la vai-se usar um nível de significância $\alpha = 5\%$ e a segunite regra de decisão: rejeitar H_0 se $Qcal > Qtab(0,95,5)$ e, portanto, o ponto crítico vale $qtab = 11,07$. O valor da estatística do teste é:

$$Qcal = \frac{(188-200)^2}{200} + \frac{(216-200)^2}{200} + \cdots + \frac{(181-200)^2}{200} + \frac{(215-200)^2}{200} = 5,74.$$

Como $5,74 < 11,07$ não rejeitamos H_0. Isto é, não há evidências que o dado não seja balanceado.

O pacote R foi usado como calculadora bem como tabela estatística. veja cuidadosamente os comandos e refaça os cálculos.

```
> O=c(188,216,209,191,181,215)#####Frequências observadas
> sum(O)
[1] 1200
> E=200*c(1,1,1,1,1,1)#####Frequências observadas
> sum(E)##soma(O)=soma(E)-\'{E} bom veficar sempre!!!!!!
[1] 1200
> D=O-E
> Q=D^2/E
> sai=cbind(O,E,D,Q);sai
     O   E   D     Q
[1,] 188 200 -12 0.720 [2,] 216 200  16 1.280 [3,] 209 200   9 0.405
[4,] 191 200  -9 0.405 [5,] 181 200 -19 1.805 [6,] 215 200  15 1.125
> Qcal=sum(Q);Qcal ##Estatística do Teste de qui-quadrado de Pearson
[1] 5.74
> Qtab=qchisq(0.95,5);Qtab;round(Qtab,2)##Valor tabelado
[1] 11.07050 [1] 11.07
```

1.23 Exercícios propostos

1. Suponha que uma variável aleatória X tenha distribuição Uniforme sobre o conjunto $A = \{10, 11, \ldots, 20\}$. Pede-se:

   a. A f.p. de X.

   b. A probabilidade de X ser par

   c. A distribuição de $Y = X - 9$.

   d. A média e a variância de Y usando o item c.

2. Um grupo de 5 estudantes tem suas idades de acordo com o conjunto $A = \{14, 16, 18, 20, 22\}$. Um estudante estudante é selecionado ao acaso e sua idade Y é anotado. Responda ao que se pede.

a. Qual a distribuição de probabilidade de Y?

b. Ache uma transformação $X = g(Y) = \dfrac{Y-12}{2}$ de sorte que Y tenha distribuição uniforme padrão de parâmetro $N = 5$.

c. Calcule a média(μ) e a variância (σ^2) de Y diretamente.

d. Calcule a média e a variância de Y usando a transformação do item b.

e. Dois estudantes são selecionados aleatoriamente com reposição e suas idades anotadas (Y_1, Y_2). Qual a função de probabilidade de W,a idade média dos estudantes selecionados? Calcule a valor esperado e a variância de W? Em Amostragem dizemos que tiramos uma amostra aleatória com reposição de tamanho $n = 2$ de uma população de tamanho $N = 5$.

f. Verifique

$$E(W) = E(\bar{Y}) = \mu = E(Y)$$

e

$$Var(W) = V(\bar{Y}) = \frac{var(\sigma^2)}{n}.$$

g. Refaça o item e para o caso sem reposição. Em Amostragem dizemos que tiramos uma amostra aleatória sem reposição de tamanho $n = 2$ de uma população de tamanho $N = 5$.

h. Verifique

$$E(W) = E(\bar{Y}) = \mu = E(Y)$$

e

$$Var(W) = V(\bar{Y}) = \frac{\sigma^2}{n}\frac{N-n}{N-1}.$$

i. A amostra sorteada com reposição foi $(16, 20)$. Estime a média populacional μ através da média amostral. Estime também o toal populacional , isto é, a soma das idades do grupo de estudantes.

3. Um estudo sobre acidentes de trabalho numa industria revelou que, em 150 acidentes,

Dia	Segunda	Terça	Quarta	Quinta	Sexta	Total
Freq. Observada	46	26	18	16	44	150

o objetivo é testar que os acidentes ocorrem com igual frequência nos cinco dias da semana. Formule a hipótese nula em termos de uma distribuição de probabilidade. Use um nível de significância de 5%. Faça comentátios.

4. Seja $X_n \sim Unif(n), n = 1, 2, \ldots$ e $Y_n = \dfrac{X_n}{n}$. Mostre que:

a. a f.g.m. de Y_n é $M_{Y_n}(t) = \dfrac{e^t - 1}{n(1 - e^{-t/n})}, \quad t \neq 0.$

b. $\lim_{n\to\infty} n(1 - e^{-t/n}) = t$

c. $\lim_{n\to\infty} M_{Y_n}(t) = \dfrac{e^t - 1}{t}, \quad t \neq 0 = M_Y(t).$

d. Usando o apêndice B do Mood que a função geradora de momentos obtida no item b é a função geradora de momentos da uniforme contínua em $(0, 1)$. Isto é uma ligação importante entre as distribuições uniforme discreta e contínua.

5. Usando a função geradora de probabilidades de $X \sim Unif(5)$, calcule:

   a. O valor esperado de X;

   b. O segundo momento fatorial

   c. A variância.

6. Usando a função geradora de momentos de $X \sim unif(4)$, calcule:

   a. O valor esperado de X;

   b. O segundo momento fatorial

   c. A variância.