Variáveis Aleatórias Discretas

Prof. Leandro Chaves Rêgo

Departamento de Estatística e Matemática Aplicada - UFC

Fortaleza, 19 de janeiro de 2022

Variável Aleatória Discreta

Descreveremos os principais modelos de variáveis aleatórias discretas, dentre esses destacam-se: Bernoulli, Binomial, Poisson, Geométrica, Pascal, Hipergeométrica, Zeta e, como um modelo de uma distribuição discreta multivariada, a Multinomial. Para cada uma delas será dada a distribuição de probabilidade, $P(X = k)$, esperança, $E(X)$ e variância, $V(X)$.
Uma explicação: parâmetro da distribuição de probabilidade é a entidade sem a qual é impossível calcular probabilidades envolvendo a variável aleatória. O k, em $P(X = k)$, é um dos valores que a variável aleatória assume com probabilidade diferente de zero, isto é, é um dos valores do seu contradomínio. Se o ou os parâmetros da distribuição de probabilidade não são conhecidos, o que acontece em problemas práticos, a Estatística fornece método para estimá-los.

Uniforme Discreta de parâmetro n.

X tem uma distribuição *aleatória* com parâmetro n, onde n é um número inteiro, se $X(w) \in \{x_1, x_2, \ldots, x_n\}$ e $p(x_i) = \frac{1}{n}$, para $i \in \{1, \ldots, n\}$.
A função de probabilidade aleatória pode ser utilizada para modelar mecanismos de jogos (por exemplo, dados e moedas balanceados, cartas bem embaralhadas). Utilizando a propriedade de aditividade da probabilidade, é fácil ver que para qualquer evento $A \subseteq \{x_1, x_2, \ldots, x_n\}$, temos que $P(X \in A) = \frac{||A||}{n}$.

Bernoulli de parâmetro p: $B(p)$

A modelagem de uma situação do mundo físico por uma Bernoulli envolve definir um evento de interesse, A por exemplo, e a ele associar uma probabilidade $p = P(A)$. Portanto, nesta modelagem, o mundo real é dicotômico, isto é, ou A acontece, ou não, neste último caso, acontece seu complementar. Assim, uma Bernoulli pode ser adequada para modelar: o estado de uma impressora, se funcionando ou não; em uma palavra de máquina um dado *bit* ser 1 ou 0. Além de modelar o mundo real, a Bernoulli é básica em desenvolvimentos teóricos como, para conjuntamente com a desigualdade de Tchebychev, provar a Lei dos Grandes Números.

Bernoulli de parâmetro p: $B(p)$

(i) Distribuição de probabilidade.

$$P(X = k) = \begin{cases} q, & k = 0, \\ p, & k = 1, \end{cases}$$

onde $q = 1 - p$ e portanto $\sum_{k=0}^{1} P(X = k) = 1$.

(ii) Esperança.

$$E(X) = 0 \times q + 1 \times p = p.$$

(iii) Variância.

$$E(X^2) = 0^2 \times p + 1^2 \times p = p,$$

logo

$$V(X) = E(X^2) - E(X)^2 = p - p^2 = pq.$$

Exemplo

Exemplo

Suponha uma única posição de memória num computador digital. O objetivo é modelar a distribuição da variável aleatória, X, que conta o número de uns (1's) nesta posição de memória. Imagine que não haja razão para supor que o aparecimento de 0 ou 1 tenham probabilidades diferentes, portanto, assume-se que $P(0) = P(1) = \frac{1}{2}$. Então, $P(X = 0) = P(X = 1) = \frac{1}{2}$. ■

Binomial de parâmetros n e p: $B(n,p)$

Uma variável binomial conta o número de ocorrências (ou o número de sucessos) de um dado evento A em n realizações de experimentos independentes de Bernoulli, onde $P(A) = p$ permanece constante em todo o desenvolvimento do experimento. Assim, uma binomial é adequada para modelar, entre outros, o número de zeros em uma palavra de máquina de precisão simples ou dupla, o número de processadores em funcionamento em um sistema multiprocessador ou o número de servidores ativos em um dado sistema de computação.

Binomial de parâmetros n e p: $B(n,p)$

(i) Distribuição de probabilidade.

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\ k=0,\ldots,n.$$

Note que, usando o *teorema binomial*[1] tem-se que

$$\sum_{k=0}^{n}P(X=k)=\sum_{k=0}^{n}\binom{n}{k}p^kq^{n-k}=(p+q)^n=1.$$

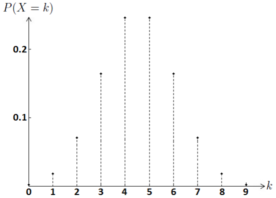

Figura: Gráfico de uma distribuição Binomial

[1] $(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}$.

Binomial de parâmetros n e p: $B(n, p)$

(ii) Esperança.

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

(iii) Variância.

Um cálculo similar ao de $E(X)$ mostra que

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

e portanto,

$$V(X) = npq + n^2 p^2 - n^2 p^2 = npq.$$

Binomial de parâmetros n e p: $B(n, p)$

Uma variável aleatória relacionada com uma $X \sim B(n, p)$ é $Y = n - X$. Neste caso, Y conta o número de falhas[2] nas n repetições independentes do experimento, sendo então uma $B(n, q)$.
Os comandos do R dbinom, pbinom, qbinom e rbinom nos fornecem valores para a função de probabilidade, função de distribuição acumulada, quantis da distribuição e amostras aleatórias da distribuição binomial, respectivamente.

[2] Na verdade a definição do que sucesso ou falha depende de como a modelagem está sendo realizada.

Exemplo

Um sistema de computação on-line tem 20 linhas de comunicação que operam independentemente. A probabilidade que qualquer linha esteja ocupada é 0.6. Qual é a probabilidade que 10 ou mais linhas estejam em operação?

Exemplo

Solução: $A = \{\text{a linha está ocupada}\}$. $P(A) = p = 0.6 \Rightarrow q = 1 - p = 0.4$.
Seja $X=$ número de linhas ocupadas, ou em operação. Logo,

$$P(X = k) = \binom{20}{k} 0.6^k 0.4^{20-k}, \; k = 0, \ldots, 20.$$

Portanto,

$$P(X \geq 10) = \sum_{k=10}^{20} \binom{20}{k} 0.6^k 0.4^{20-k}.$$

$\blacksquare$

Poisson de parâmetro λ: $Poisson(\lambda)$

A função de probabilidade Poisson é usualmente utilizada para modelar a contagem do número de ocorrências de eventos aleatórios em um certo tempo t, como por exemplo o número de fótons emitidos por uma fonte de luz de intensidade I fótons/seg em t segundos ($\lambda = It$), o número de clientes chegando em uma fila no tempo t ($\lambda = Ct$), o número de ocorrências de eventos raros no tempo t ($\lambda = Ct$), entre outros.

Poisson de parâmetro λ: *Poisson*(λ)

(i) Distribuição de probabilidade.

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, k \in \{0, 1, \ldots\}.$$

Usando o resultado da expansão em série de Taylor da função exponencial, sabe-se que para todo x real,

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Portanto,

$$\sum_{k=0}^{\infty} P(X = k) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

Poisson de parâmetro λ: $Poisson(\lambda)$

(ii) Esperança.

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!} = \sum_{k=1}^{\infty} \frac{e^{-\lambda}\lambda^k}{(k-1)!} = \sum_{s=0}^{\infty} \frac{e^{-\lambda}\lambda^{s+1}}{s!} \\ &= \lambda e^{-\lambda} \sum_{s=0}^{\infty} \frac{e^{-\lambda}\lambda^s}{s!} = \lambda e^{-\lambda} e^{\lambda} = \lambda. \end{aligned}$$

No cálculo anterior, $k - 1 = s \Rightarrow k = s + 1$.

Poisson de parâmetro λ: *Poisson*(λ)

(iii) Variância.

$$\begin{aligned} E(X^2) &= \sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda}\lambda^k}{k!} = \sum_{k=1}^{\infty} k \frac{e^{-\lambda}\lambda^k}{(k-1)!} = \sum_{s=0}^{\infty}(s+1)\frac{e^{-\lambda}\lambda^{s+1}}{s!} \\ &= \sum_{s=0}^{\infty} s \frac{e^{-\lambda}\lambda^{s+1}}{s!} + \sum_{s=0}^{\infty} \frac{e^{-\lambda}\lambda^{s+1}}{s!} \\ &= \lambda \sum_{s=0}^{\infty} s \frac{e^{-\lambda}\lambda^{s}}{s!} + \lambda \sum_{s=0}^{\infty} \frac{e^{-\lambda}\lambda^{s}}{s!} \\ &= \lambda^2 + \lambda. \end{aligned}$$

Portanto,

$$V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Os comandos do R dpois, ppois, qpois e rpois nos fornecem valores para a função de probabilidade, função de distribuição acumulada, quantis da distribuição e amostras aleatórias da distribuição Poisson, respectivamente.

Exemplo

Se a probabilidade de 0 fótons serem emitidos no tempo $[0, t)$ é igual a 0.1, então qual a probabilidade de que pelo menos 2 fótons serem emitidos no mesmo tempo t?

Exemplo

Solução: $A = \{\text{emissão de fótons no tempo } t\}$. Seja $X=$ número de fótons serem emitidos no tempo t. Neste problema não foi dado o parâmetro λ, mas foi fornecida uma condição para encontrá-lo.

$$P(X = 0) = 0.1 \Rightarrow \frac{e^{-\lambda}\lambda^0}{0!} = 0.1 \Rightarrow e^{-\lambda} = 0.1 \Rightarrow \lambda = 2.302.$$

Portanto, $P(X \geq 2) = 1 - P(X \leq 1) = 1 - e^{-2.302} - 2.302e^{-2.302}$. ■

Moda da Poisson

Exemplo

Um valor mais provável de uma distribuição de Poisson é k^* se

$$P(X = k^*) \geq P(X = k^* + 1)$$

e

$$P(X = k^*) \geq P(X = k^* - 1).$$

Realizando estes cálculos, estas condições são equivalentes a,

$$k^* \leq \lambda \leq k^* + 1$$

ou

$$\lambda - 1 \leq k^* \leq \lambda.$$

Se k^* é o maior inteiro menor ou igual a λ esta restrição é satisfeita, e portanto este é um valor mais provável desta distribuição. Em outras palavras, k^* é o valor de k que torna máxima a probabilidade na Poisson. ■

Poisson como um Limite de Eventos Raros de Binomial

A distribuição de Poisson pode ser encontrada pelo limite de uma $B(n,p)$, quando n vai para infinito (isto é, o experimento é realizado um número grande de vezes), p é muito pequeno mas np, que é média da binomial, permanece constante. A explanação a seguir[3] motiva como essa aproximação pode ser realizada.

Suponha que chamadas telefônicas cheguem em uma grande central e que em um período particular de três horas (180 minutos) um total de 270 chamadas tenham sido recebidas, ou seja, 1.5 chamadas por minuto. O objetivo é calcular a probabilidade de serem recebidas k chamadas durante os próximos três minutos.

[3] de P. Meyer (1983) pág. 187

Poisson como um Limite de Eventos Raros de Binomial

É natural pensar que a qualquer instante pode ocorrer uma chamada, portanto a modelagem do problema exige que aproximações sejam feitas.
Para começar, pode-se dividir o intervalo de 3 minutos em nove intervalos de 20 segundos cada um e tratar cada um desses nove intervalos como um ensaio de Bernoulli, durante o qual observa-se uma chamada (sucesso) ou nenhuma chamada (falha), com probabilidade de sucesso igual a $p = 1.5 \times \frac{20}{60} = 0.5$.
Desse modo, a tentação é grande para afirmar que a probabilidade de 2 chamadas é igual a $\binom{9}{2}(0.5)^9 = \frac{9}{128}$.

Poisson como um Limite de Eventos Raros de Binomial

Porém, este cálculo ignora a possibilidade de que mais de uma chamada possa ocorrer em um único intervalo. Então, por que não aumentar o número n de subintervalos de tempo de modo que cada subintervalo corresponda a $\frac{180}{n}$ segundos e portanto a probabilidade de ocorrência de uma chamada em um subintervalo seja igual a $p = 1.5 \times \frac{180}{60n}$? Desta maneira $np = 4.5$ permanece constante quando o número de subintervalos cresce. Utilizando novamente o modelo binomial, a probabilidade de ocorrerem k chamadas é dada por: $\binom{n}{k}(\frac{4.5}{n})^k(1 - \frac{4.5}{n})^{n-k}$. O que acontece com esta probabilidade quando $n \to \infty$?

Poisson como um Limite de Eventos Raros de Binomial

A resposta, como será visto a seguir, é que esta distribuição tende a uma distribuição de Poisson, sendo este resultado conhecido como *limite de eventos raros*.
Seja a expressão geral da probabilidade binomial,

$$\begin{aligned} P(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{k!} p^k (1-p)^{n-k}. \end{aligned}$$

Como o objetivo é estudar o caso em que np é constante, seja $np = \lambda$, ou seja, $p = \frac{\lambda}{n}$ e $1 - p = \frac{n-\lambda}{n}$. Então,

Poisson como um Limite de Eventos Raros de Binomial

$$
\begin{aligned}
P(X=k) &= \frac{n(n-1)\cdots(n-k+1)}{k!}(\frac{\lambda}{n})^k(\frac{n-\lambda}{n})^{n-k} \\
&= \frac{\lambda^k}{k!}((1)(1-\frac{1}{n})\cdots(1-\frac{k-1}{n})(1-\frac{\lambda}{n})^{n-k}.
\end{aligned}
$$

Fazendo $n \to \infty$, os termos da forma $(1-\frac{j}{n})$, para $1 \leq j \leq k-1$, tendem para 1 e como existe um número fixo $k-1$ deles, o seu produto também tende a 1. O mesmo ocorre com $(1-\frac{\lambda}{n})^{-k}$. Finalmente, por definição do número e, tem-se que $(1-\frac{\lambda}{n})^n \to e^{-\lambda}$ quando $n \to \infty$. Portanto,

$$\lim_{n\to\infty, p\to 0, \lambda=np} P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!},$$

ou seja, obteve-se a expressão de Poisson.

Exemplo

Ao formar números binários com n dígitos, a probabilidade de que um dígito incorreto possa aparecer é 0.002. Se os erros forem independentes, qual é a probabilidade de encontrar k dígitos incorretos em um número binário de 25 dígitos? Se um computador forma 10^6 desses números de 25 dígitos por segundo, qual é a probabilidade de que pelo menos um número incorreto seja formado durante qualquer período de 1 segundo?

Exemplo

Solução: A probabilidade de que k dígitos sejam incorretos em um número binários de 25 dígitos é igual a $\binom{25}{k}(0.002)^k(0.998)^{25-k}$. Em particular, a probabilidade de que pelo menos um dígito seja incorreto é igual a $1-(0.998)^{25} \approx 0.049$. Usando a aproximação pela Poisson então tem-se uma Poisson com parâmetro $25 \times 0.002 = 0.05$, logo a probabilidade de pelos menos um dígito incorreto neste número de 25 dígitos é $1 - e^{-0.05} \approx 0.049$.
A probabilidade de que pelo menos um número incorreto seja formado durante um período de 1 segundo é igual a $1-(0.049)^{10^6} \approx 1 - e^{-49000} \approx 1$. ■

Exemplo

A distribuição de Poisson também pode ser deduzida de específicas hipóteses relativas a determinada classe de fenômenos estudados sendo, neste caso, o parâmetro λ obtido proporcionalmente.
Suponha que o número de clientes que chegam em um banco segue uma distribuição de Poisson. Se a probabilidade de chegarem 3 clientes for o triplo da de chegarem 4 clientes em um dado período de 10 minutos. Determine:

(a) Qual o número esperado de clientes que chegam em um período de 1 hora neste banco?

(b) Qual o número mais provável de clientes que chegam em um período de 1 hora neste banco?

Exemplo

Solução: Seja X_{10} = número de clientes que chegam ao banco em 10 minutos.

$$P(X_{10} = 3) = 3P(X_{10} = 4) \Rightarrow \frac{e^{-\lambda}\lambda^3}{3!} = 3 \times \frac{e^{-\lambda}\lambda^4}{4!} \Rightarrow \lambda = 1.333.$$

(a) Seja X_{60} = número de clientes que chegam ao banco em 60 minutos. Para $t = 10$ minutos tem-se que $\lambda = 4/3$, logo para $t = 60$ minutos, $\lambda = \frac{60 \times 4/3}{10} = 8$ e portanto $E(X_{60}) = 8$.

(b) $7 \leq x^* \leq 8 \Rightarrow x^* = 7$ ou $x^* = 8$, de modo que estes são os valores mais prováveis do número de clientes que chegam no banco em 1 hora.

■

Geométrica de parâmetro p: $G(p)$

A geométrica pode ser utilizada para modelar o número de repetições independentes do lançamento de uma moeda até a primeira ocorrência de cara, tempo de espera medido em unidades de tempo inteiras até a chegada do próximo consumidor em uma fila, ou até a próxima emissão de um fóton. Esta variável, assim como as anteriores, também é uma variável de contagem, só que ela está relacionada à *primeira* ocorrência de sucesso do evento A de interesse na modelagem. Por exemplo, se o evento de interesse é a ocorrência do primeiro 1 numa *string* de zeros e uns, se a variável assumir o valor 10, a string observada foi 0000000001.

Geométrica de parâmetro p: $G(p)$

(i) Distribuição de probabilidade.

$$P(X = k) = q^{k-1}p, k \in \{1, 2, 3, \ldots\},$$

porque $\{X = k\}$ é equivalente $\underbrace{\overline{A} \cap \overline{A} \cap \ldots \cap \overline{A}}_{k-1} \cap A$.

Utilizando o resultado de uma soma infinita de uma progressão geométrica ilimitada de razão $| r | < 1$,

$$\sum_{k=1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} q^{k-1}p = p \sum_{k=1}^{\infty} q^{k-1} = 1.$$

Logo, esta é uma legítima função probabilidade de massa.

Geométrica de parâmetro p: $G(p)$

(ii) Esperança.

$$\begin{aligned}
E(X) &= \sum_{k=1}^{\infty} kP(X=k) = \sum_{k=1}^{\infty} kq^{k-1}p \\
&= p\sum_{k=1}^{\infty} kq^{k-1} = p\sum_{k=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}q}q^{k} \\
&= p\frac{\mathrm{d}}{\mathrm{d}q}\sum_{k=1}^{\infty} q^{k} = p\frac{\mathrm{d}}{\mathrm{d}q}\left(\frac{q}{1-q}\right) = \frac{1}{p}.
\end{aligned}$$

(iii) Função Geradora de Momentos.

$$\begin{aligned}
M_X(t) &= \sum_{k=1}^{\infty} q^{k-1}pe^{tk} = \frac{p}{q}\sum_{k=1}^{\infty}(qe^{t})^{k} \\
&= \frac{p}{q}\frac{qe^{t}}{1-qe^{t}} = \frac{pe^{t}}{1-qe^{t}}.
\end{aligned}$$

Geométrica de parâmetro p: $G(p)$

(iv) Variância. Usando a função geradora de momentos tem-se que

$$E(X^2) = M_X^{(2)}(0).$$

Derivando duas vezes $M_X(t)$, temos

$$M_X^{(1)}(t) = \frac{(1-qe^t)pe^t - pe^t(-qe^t)}{(1-qe^t)^2} = \frac{pe^t}{(1-qe^t)^2}.$$

e

$$M_X^{(2)}(t) = \frac{(1-qe^t)^2 pe^t - pe^t(2(1-qe^t))(-qe^t)}{(1-qe^t)^4}.$$

Logo,

$$E(X^2) = \frac{(1-q)^2 p + p(2p)q}{p^4} = \frac{1+q}{p^2}.$$

e

$$V(X) = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}.$$

Geométrica no R

Os comandos do R dgeom, pgeom, qgeom e rgeom nos fornecem valores para a função de probabilidade, função de distribuição acumulada, quantis da distribuição e amostras aleatórias da distribuição Geométrica, respectivamente.

Exemplo

Suponha que joga-se uma moeda independentemente até que uma coroa ocorra. Sabe-se que probabilidade de cara igual a $0 < p < 1$. Seja X o número de repetições necessárias até que coroa apareça pela primeira vez na sequência. Qual é a probabilidade do evento $\{X = k\}$ para $k \in \{1, 2, 3, \dots\}$? Note que para que $X = k$ é necessário que os primeiros $k - 1$ lançamentos sejam caras e o k-ésimo lançamento seja coroa, logo, pela independência dos lançamentos, $P(X = k) = p^{k-1}q$. Ou seja, X é uma variável geométrica de parâmetro q. ■

Exemplo

Suponha que X tenha uma distribuição geométrica com parâmetro β. Mostre que para quaisquer dois inteiros positivos s e t,

$$P(X > s + t | X > s) = P(X > t).$$

Exemplo

Solução:

$$P(X > s + t | X > s) = \frac{P(X > s + t, X > s)}{P(X > s)} = \frac{P(X > s + t)}{P(X > s)}.$$

Mas

$$P(X > s + t) = \sum_{k=s+t+1}^{\infty} (1 - \beta)^{k-1}\beta = (1 - \beta)^{s+t}.$$

Similarmente, $P(X > s) = (1 - \beta)^s$ e $P(X > t) = (1 - \beta)^t$. Portanto,

$$P(X > s + t | X > s) = \frac{(1 - \beta)^{s+t}}{(1 - \beta)^s} = (1 - \beta)^t = (X > t).$$

Esta propriedade da distribuição geométrica é conhecida como *falta de memória*[4]. ■

[4] Como será visto no curso de Probabilidade 2, a variável Exponencial também tem essa propriedade.

Pascal de parâmetros r e p: $Pascal(p, r)$

Esta distribuição pode ser considerada como uma generalização da distribuição geométrica. Suponha que o interesse seja calcular a probabilidade de que um experimento tenha de ser repetido k vezes para que um evento A ocorra r vezes. Seja X o número de repetições necessárias a fim de que um evento A possa ocorrer exatamente r vezes. $X = k$ se, e somente se, A ocorrer na k-ésima repetição e A tiver ocorrido $r - 1$ vezes nas $(k - 1)$ repetições anteriores.

Pascal de parâmetros r e p: $Pascal(p, r)$

Uma possível realização do experimento é

$$B = \underbrace{\overline{A} \cap \ldots \cap \overline{A}}_{k-r} \cap \underbrace{A \ldots \cap A}_{r}.$$

Assumindo independência entre os eventos, a probabilidade acima corresponde a

$$\underbrace{q \times \ldots \times q}_{k-r} \times \underbrace{p \times \ldots \times p}_{r} = q^{k-r} p^{r}.$$

Mas, quantas realizações distintas do evento B são possíveis? A resposta é $\binom{k-1}{r-1}$. Portanto,

Pascal de parâmetros r e p: $Pascal(p, r)$

(i) Distribuição de probabilidade.

$$P(X = k) = \binom{k-1}{r-1} p^r q^{k-r}, \; k \geq r.$$

Se $r = 1$, tem-se que X tem uma distribuição geométrica com parâmetro p.

Pascal de parâmetros r e p: $Pascal(p, r)$

(ii) Esperança e Variância.

Para calcular $E(X)$ e $V(X)$ pode-se proceder da seguinte maneira. Seja $Z_1, Z_2, \ldots$ uma sequência de variáveis aleatórias tal que Z_1 é o número de repetições do experimento até a primeira ocorrência de um evento A, Z_i é o número de repetições do experimento entre a $(i-1)$-ésima até e incluindo a i-ésima ocorrência de A, para $i = 2, 3, \ldots, r$, isto é,

$$\underbrace{\overline{A} \cap \ldots \cap \overline{A}}_{Z_1} \cap A \cap \underbrace{\overline{A} \cap \ldots \cap \overline{A}}_{Z_2} \cap A \ldots \cap \underbrace{\overline{A} \cap \ldots \cap \overline{A}}_{Z_r} \cap A.$$

Então, as variáveis Z_i são independentes, cada uma delas tem uma distribuição geométrica com parâmetro p e $X = Z_1 + Z_2 + \cdots + Z_r$. Logo, X pode ser considerada como uma soma de r geométricas independentes, portanto, usando propriedades da esperança e da variância,

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

e

$$V(X) = \frac{r(1-p)}{p^2}.$$

Exemplo

Suponha que X tenha distribuição binomial com parâmetros n e p e Y tenha uma distribuição de Pascal com parâmetros r e p. Portanto, $P(X \geq r) = P(Y \leq n)$.
Estas duas distribuições tratam de ensaios de Bernoulli repetidos. A distribuição binomial surge quando se tem um número fixo de ensaios e o interesse é o número de sucessos que venham a ocorrer. A distribuição de Pascal é encontrada quando o número de sucessos é fixo, r, e o que é registrado é o número de ensaios necessários para a obtenção dos r sucessos. ■

Hipergeométrica de parâmetros N, D, e n: $H(N, D, n)$

A distribuição hipergeométrica descreve o número de sucessos em uma sequência de n amostras retiradas sem reposição de uma população finita. Por exemplo, considere que tem-se uma carga com N objetos dos quais D são defeituosos. A distribuição hipergeométrica descreve a probabilidade de que em uma amostra de n objetos distintos escolhidos da carga aleatoriamente exatamente k objetos sejam defeituosos.

Hipergeométrica de parâmetros N, D, e n: $H(N, D, n)$

(i) Distribuição de probabilidade.

$$P(X = k) = \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}}.$$

Esta probabilidade é positiva se: $N - D \geq n - k$, ou seja $k \geq \max(0, D + n - N)$ e $k \leq \min(n, D)$.

Esta fórmula pode ser entendida assim: existem $\binom{N}{n}$ possíveis amostras sem reposição, $\binom{D}{k}$ maneiras de escolher k objetos defeituosos e $\binom{N-D}{n-k}$ maneiras de preencher o resto da amostra com objetos sem defeito.

Quando a população é grande quando comparada ao tamanho da amostra (ou seja, N for muito maior que n) a distribuição hipergeométrica é aproximada por uma distribuição binomial com parâmetros n (tamanho da amostra) e $p = D/N$ (probabilidade de sucesso em um único ensaio) (Meyer, 1983).

Hipergeométrica de parâmetros N, D, e n: $H(N, D, n)$

(ii) Esperança.

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

Hipergeométrica de parâmetros N, D, e n: $H(N, D, n)$

Substituindo no somatório $D^* = D - 1, k^* = k - 1, n^* = n - 1$ e $N^* = N - 1$,

$$E(X) = \frac{nD}{N} \sum_{k^*=0}^{n^*} \frac{\binom{D^*}{k^*}\binom{N^*-D^*}{n^*-k^*}}{\binom{N^*}{n^*}} = \frac{nD}{N}.$$

Se $p = \frac{D}{N}$ então, $E(X) = np$.
O somatório acima é igual a soma da função probabilidade de massa de uma variável aleatória hipergeométrica para todos os valores que tem probabilidade positiva, e portanto, é igual a 1.

Hipergeométrica de parâmetros N, D, e n: $H(N, D, n)$

(iii) Variância.

$$
\begin{aligned}
E(X^2) &= \sum_{k=0}^{n} k^2 \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}} \\
&= \sum_{k=0}^{n} (k + k(k-1)) \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}} \\
&= EX + \sum_{k=0}^{n} k(k-1) \frac{D!(N-D)!(N-n)!n!}{k!(D-k)!(n-k)!(N-D-n+k)!N!} \\
&= \frac{nD}{N} + \frac{n(n-1)D(D-1)}{N(N-1)} \times \\
&\quad \times \sum_{k=2}^{n} \frac{(D-2)!(N-D)!(N-n)!(n-2)!}{(k-2)!(D-k)!(n-k)!(N-D-n+k)!(N-2)!} \\
&= \frac{nD}{N} + \frac{n(n-1)D(D-1)}{N(N-1)} \sum_{k=2}^{n} \frac{\binom{D-2}{k-2}\binom{N-D}{n-k}}{\binom{N-2}{n-2}} \\
&= \frac{nD}{N} + \frac{n(n-1)D(D-1)}{N(N-1)} = \frac{nD}{N}\left(1 + \frac{(n-1)(D-1)}{N-1}\right),
\end{aligned}
$$

Hipergeométrica de parâmetros N, D, e n: $H(N, D, n)$

A penúltima igualdade no desenvolvimento anterior decorre do somatório consistir da soma de todas as probabilidade de uma v.a. hipergeométrica com parâmetros: $D^* = D - 2, k^* = k - 2, n^* = n - 2$ e $N^* = N - 2$.
Logo,

$$
\begin{aligned}
V(X) &= D \times \frac{n}{N}(1 + \frac{(D-1)(n-1)}{N-1}) - D^2 \times \frac{n^2}{N^2} \\
&= D \times \frac{n}{N}(\frac{(D-1)(n-1)}{N-1} + 1 - D \times \frac{n}{N}) \\
&= n \times \frac{D}{N} \times \frac{N-n}{N-1}(1 - \frac{D}{N}).
\end{aligned}
$$

Fazendo a mesma substituição no caso de $E(X)$ tem-se que:

$$V(X) = np(1-p)\frac{N-n}{N-1} = npq\frac{N-n}{N-1}.$$

Exemplo

Suponha que uma urna contêm 20 bolas brancas e 10 bolas pretas. Se 4 bolas são retiradas da urna. Determine:

(a) A probabilidade de pelo menos uma bola seja branca, se as bolas são retiradas com reposição.

(b) A probabilidade de pelo menos uma bola seja branca, se as bolas são retiradas sucessivamente e sem reposição.

(c) A probabilidade de pelo menos uma bola seja branca, se as bolas são retiradas simultaneamente.

Exemplo

Solução: Seja $B_i = \{i\text{-ésima bola é branca}, i = 1, 2, 3, 4.\}$

(a)

$$\begin{aligned} P(B_1 \cup \cdots \cup B_4) &= 1 - P(\overline{B_1 \cup \cdots \cup B_4}) \\ &= 1 - P(\overline{B_1} \cap \cdots \cap \overline{B_4}) \\ &= 1 - P(\overline{B_1}) \times \cdots \times P(\overline{B_4}) \\ &= 1 - (\frac{10}{30})^4. \end{aligned}$$

Este item também poderia ser resolvido usando o Teorema da Inclusão-Exclusão, mas daria muito trabalho. Observe que foi usado o fato dos eventos serem independentes.

(b)

$$\begin{aligned} & P(B_1 \cup \cdots \cup B_4) = 1 - P(\overline{B_1 \cup \cdots \cup B_4}) \\ = \; & 1 - P(\overline{B_1} \cap \cdots \cap \overline{B_4}) \\ = \; & 1 - P(\overline{B_1})P(\overline{B_2}|\overline{B_1}) \times \cdots \times P(\overline{B_4}|\overline{B_1} \cap \cdots \cap \overline{B_4}) \\ = \; & 1 - \frac{10}{30}\frac{9}{29}\frac{8}{28}\frac{7}{27}. \end{aligned}$$

Exemplo

(c) Seja $X = \{$número de bolas brancas retiradas da urna.$\}$ Logo

$$P(X = k) = \frac{\binom{20}{k}\binom{10}{4-k}}{\binom{30}{4}}, \; k = 0, \ldots, 4.$$

Portanto, $P(X \geq 1) = \sum_{k=1}^{4} \frac{\binom{20}{k}\binom{10}{4-k}}{\binom{30}{4}}$.

■

Exemplo

Por engano 3 peças defeituosas foram misturadas com boas formando um lote com 12 peças no total. Escolhendo ao acaso 4 dessas peças, determine a probabilidade de encontrar:

(a) Pelo menos 2 defeituosas.

(b) No máximo 1 defeituosa.

(c) No mínimo 1 boa.

Exemplo

Solução: Seja $X = \{$número de peças defeituosas no lote.$\}$ Logo

$$P(X = k) = \frac{\binom{3}{k}\binom{9}{4-k}}{\binom{12}{4}}, \ k = 0, \ldots, 4.$$

(a) $P(X \geq 2) = P(X = 2) + P(X = 3) = \frac{\binom{3}{2}\binom{9}{2}}{\binom{12}{4}} + \frac{\binom{3}{3}\binom{9}{1}}{\binom{12}{4}}.$

(b) $P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{\binom{3}{0}\binom{9}{4}}{\binom{12}{4}} + \frac{\binom{3}{1}\binom{9}{3}}{\binom{12}{4}}.$

(c) Seja $Y = \{$número de peças perfeitas na amostra.$\}$. O contradomínio de Y é $R_Y = \{1, 2, 3, 4\}$ e portanto $P(Y \geq 1) = 1.$

■

Zeta ou Zipf de parâmetro $\alpha > 1$: $Z(\alpha)$

A função probabilidade Zeta ou Zipf é um exemplo de uma distribuição de caudas-pesadas[5] cuja importância cresceu desde meados dos anos 1990. As aplicações desta função de probabilidade incluem: número de consumidores afetados por um *blackout*, tamanhos de arquivos solicitados em transferência via *Web* e atraso de pacotes na *internet*.

[5] Se uma variável aleatória tem caudas-pesadas então tem probabilidades grandes nas caudas.

Zeta ou Zipf de parâmetro $\alpha > 1$: $Z(\alpha)$

(i) Distribuição de probabilidade.

$$P(X = k) = \frac{k^{-\alpha}}{\zeta(\alpha)}, k = 1, 2, \dots, \quad \alpha > 1$$

onde $\zeta(\alpha) = \sum_{j=1}^{\infty} j^{-\alpha}$ é conhecida como a função Zeta de Riemann, sendo convergente para $\alpha > 1$.

(ii) Esperança.

$$E(X) = \sum_{k=1}^{\infty} k \frac{k^{-\alpha}}{\zeta(\alpha)} = \frac{1}{\zeta(\alpha)} \sum_{k=1}^{\infty} k^{-(\alpha-1)} = \frac{\zeta(\alpha-1)}{\zeta(\alpha)}, \ \alpha > 2.$$

(iii) Variância.

$$E(X^2) = \sum_{k=1}^{\infty} k^2 \frac{k^{-\alpha}}{\zeta(\alpha)} = \frac{1}{\zeta(\alpha)} \sum_{k=1}^{\infty} k^{-(\alpha-2)} = \frac{1}{\zeta(\alpha)} \zeta(\alpha-2), \alpha > 3.$$

Logo,

$$V(X) = \frac{1}{\zeta(\alpha)} \zeta(\alpha - 2) - (\frac{\zeta(\alpha-1)}{\zeta(\alpha)})^2.$$

Exemplo

Os tamanhos de arquivos armazenados em um grande sistema de arquivos Unix seguem uma distribuição Zeta com parâmetro α quando estes tamanhos são medidos em kilobytes[6].

(a) Se os tamanhos dos arquivos de 1KB são 10000 vezes mais prováveis que tamanhos de arquivos de 1MB, então qual o valor do parâmetro α?

(b) Quanto mais provável são tamanhos de arquivos de 1MB em comparação com tamanhos de arquivos de 1GB?

[6] 1024bytes = 1kilobyte = 1KB; 1024kilobytes = 1megabyte = 1MB; 1024megabytes = 1gigabytes = 1GB; 1024gigabytes = 1terabyte = 1TB; 1024terabytes = 1petabyte = 1PB; 1024petabytes = 1exabyte = 1EB; 1024exabytes = 1zettabyte = 1ZB; 1024zettabytes = 1yottabyte = 1YB.

Exemplo

Solução:

(a) Seja $X = \{\text{tamanho dos arquivos em KB}\}$. Logo,

$$\begin{aligned} P(X=1) &= 10000P(X=1024) \Rightarrow \frac{1}{\zeta(\alpha)} = 10000\frac{1024^{-\alpha}}{\zeta(\alpha)} \\ &\Rightarrow \alpha = \frac{log 10.000}{log 1024} = 1.329. \end{aligned}$$

(b) $P(X=1024) = \frac{1024^{-\alpha}}{\zeta(\alpha)}$ e $P(X=1024^2) = \frac{(1024^2)^{-\alpha}}{\zeta(\alpha)}$.
Portanto,

$$\frac{P(X=1024)}{P(X=1024^2)} = 1024^{\alpha} = 10.000.$$

■

Multinomial

A Multinomial é uma distribuição conjunta de variáveis aleatórias discretas, que pode ser considerada como uma generalização da distribuição binomial. Considere um experimento aleatório qualquer e suponha que o espaço amostral deste experimento é particionado em k eventos $A_1, A_2, \ldots, A_k$, onde o evento A_i tem probabilidade p_i. Suponha que se repita este experimento n vezes de maneira independente e seja X_i o número de vezes que o evento A_i ocorreu nestas n repetições. Então,

$$P(X_1 = n_1, X_2 = n_2, \ldots, X_k = n_k) = \frac{n!}{n_1! n_2! \cdots n_k!} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k},$$

onde $\sum_{i=1}^{k} n_i = n$.
Lembrando que o número de maneiras de arranjar n objetos, n_1 dos quais é de uma espécie, n_2 dos quais é de uma segunda espécie, ..., n_k dos quais são de uma k-ésima espécie é dado pelo coeficiente multinomial $\frac{n!}{n_1! n_2! \cdots n_k!}$, que corresponde a uma permutação com repetição.

Hipergeométrica Multivariada

A Hipergeométrica Multivariada é uma distribuição conjunta de variáveis aleatórias discretas, que pode ser considerada como uma generalização da distribuição Hipergeométrica.
Considere que um lote de tamanho N com peças de k tipos diferentes.
Suponha que existam N_i peças do tipo i no lote, $i = 1, 2, \ldots, k$, que n peças são sorteadas aleatoriamente do lote, sem reposição, e que X_i seja o número de peças do tipo i na amostra retirada do lote. Então,

$$P(X_1 = n_1, X_2 = n_2, \ldots, X_k = n_k) = \frac{\prod_{i=1}^{k} \binom{N_i}{n_i}}{\binom{N}{n}},$$

em que $\sum_{i=1}^{k} n_i = n$ e $\sum_{i=1}^{k} N_i = N$.

Exemplo

Uma impressora está apresentando uma média de 2 erros por página impressa. Sabe-se que não há razão para supor que exista qualquer relação entre os erros (ou acertos) de uma página para outra, e que o fenômeno obedece a uma lei de Poisson.

(a) Qual é a probabilidade de que uma página não apresente erros?

(b) Uma amostra aleatória de 10 páginas é selecionada de todas as páginas impressas pela impressora. Qual é a probabilidade de que essa amostra apresente 2 ou mais páginas sem erro?

(c) Qual é a probabilidade de que a primeira página sem erro apareça na décima impressão?

Exemplo

(d) Suponha que páginas são impressas até que ocorram 3 sem erro. Qual é a probabilidade de que seja necessário imprimir 5 páginas para que isto aconteça?

(e) Cada página impressa é classificada como *excelente* se não apresentar erros, *boa*, se contiver de 1 a 2 erros e *inaceitável*, se contiver 3 ou mais erros. Se 15 páginas são inspecionadas, qual é a probabilidade de que 8, 5 e 2 páginas, respectivamente, pertencam a cada uma das categorias acima?

Exemplo

Solução:

(a) Seja $N_p = \{\text{número de erros por página.}\}$. Logo,

$$N_p \sim Poisson(\lambda = 2)$$

e a probabilidade pedida é

$$P(N_p = 0) = e^{-2} = 0.136.$$

(b) Seja $N = \{\text{número de páginas sem erro.}\}$. Logo,

$$N \sim B(n = 10, p = P(N_p = 0) = 0.136).$$

Portanto,

$$P(N \geq 2) = 1 - P(N \leq 1) = 1 - \binom{10}{0} p^0 q^{10} - \binom{10}{1} p^1 q^9.$$

(c) $N_1 =$
$\{\text{número de páginas impressas até aparecer a primeira sem erro.}\}$

$$N_1 \sim G(p = P(N_p = 0))$$

e assim,

$$P(N_1 = 10) = (1 - e^{-2})^9 e^{-2}.$$

Exemplo

(d) $N_2 = \{$ número de páginas que tem de ser impressas para aparecerem 3 sem erro.$\}$,

$$N_2 \sim Pascal(p = P(N_p = 0), r = 3).$$

Logo,

$$P(N_2 = 5) = \binom{5-1}{3-1}(e^{-2})^3(1 - e^{-2})^{5-3}.$$

(e) A distribuição aqui é uma Multinomial. Seja X_i, $i = 1, 2, 3$, número de páginas classificadas como *excelente*, *boa* e *inaceitável*, respectivamente. Portanto,

$$P(X_1 = 8, X_2 = 5, X_3 = 2) = \frac{15!}{8!5!2!} p_1^8 p_2^5 p_3^2,$$

onde

$$p_1 = P(N_p = 0),\ \ p_2 = P(1 \le N_p \le 2),\ \ p_3 = P(N_p \ge 3).$$

Exemplo

Em determinado cruzamento da BR101 caminhões transitam exibindo placas de vários estados; estima-se que em um dado intervalo de tempo, t, as probabilidades de transitarem neste cruzamento caminhões com placas do Ceará (CE), Amazonas (AM) e Maranhão (MA), são, respectivamente, $\frac{1}{25}$, $\frac{1}{100}$ e $\frac{1}{80}$, sendo $\frac{15}{16}$ a probabilidade de caminhões com placas que não sejam destes três estados.

(a) Se no intervalo de tempo dado passarem neste cruzamento 9 caminhões, qual é a probabilidade de que, dentre os 9, um igual número deles proceda somente de cada um dos referidos estados?

(b) Sabe-se que o número de caminhões com placa do Ceará, que transitam neste cruzamento e no dado intervalo de tempo, segue uma distribuição de Poisson com taxa (ou média) $\frac{1}{25}$ caminhões por dia. Qual é a probabilidade de que em uma semana, pelo menos um caminhão apresente placa do Ceará?

Exemplo

(c) Sabe-se que 2004 terá 53 semanas. Qual é a probabilidade de que o número de semanas onde trafega pelo menos um caminhão com placa do Ceará seja exatamente 2?

(d) Qual é a probabilidade de que a décima semana seja a primeira onde transita pelo menos um caminhão com placa do Ceará?

(e) Qual é a probabilidade de que seja necessário transcorrerem 20 semanas a fim de que em exatamente 3 delas pelos menos um caminhão apresente placa do Ceará?

Exemplo

Solução: Sejam os eventos $C = \{$ a placa é do Ceará$\}$, $A = \{$ a placa é do Amazonas$\}$, $M = \{$ a placa é do Maranhão$\}$ e $N = \{$ a placa não é do CE, AM e MA$\}$. Portanto, $P(C) = \frac{1}{25}$, $P(A) = \frac{1}{100}$, $P(M) = \frac{1}{80}$ e $P(N) = \frac{15}{16}$.

(a) Distribuição: Multinomial. Seja X_i, $i = 1, 2, 3, 4$, respectivamente, número de caminhões com placas do CE, AM, MA e com placas que não sejam destes três estados.

$$
\begin{aligned}
& P(X_1 = 3, X_2 = 3, X_3 = 3, X_4 = 0) \\
= \quad & \frac{9!}{3!3!3!0!}\left(\frac{1}{25}\right)^3\left(\frac{1}{100}\right)^3\left(\frac{1}{80}\right)^3\left(\frac{15}{16}\right)^0.
\end{aligned}
$$

Exemplo

(b) $X_1 \sim Poisson(\lambda = \frac{1}{25})$. Portanto, se em 1 dia, $\lambda = \frac{1}{25}$, em uma semana, isto é, 7 dias, $\lambda = \frac{7}{25}$.. Seja X_s={número de caminhões com placa do Ceará trafegando em uma semana}. Logo,

$$X_s \sim Poisson(\lambda = \frac{7}{25}),$$

e

$$P(X_s \geq 1) = 1 - P(X_s < 1) = 1 - e^{-\frac{7}{25}}.$$

Exemplo

(c) Seja $X_C = \{$número de semanas onde trafega pelo menos um caminhão com placa do Ceará$\}$. Então

$$X_C \sim B(n = 53, p = 1 - e^{-\frac{7}{25}})$$

e

$$P(X_C = 2) = \binom{53}{2} p^2 (1-p)^{51}.$$

(d) Distribuição Geométrica.

$$P(X_C = 10) = (e^{-\frac{7}{25}})^9 (1 - e^{-\frac{7}{25}}).$$

(e) Distribuição de Pascal.

$$P(X_C = 3) = \binom{19}{2} (1 - e^{-\frac{7}{25}})^3 (e^{-\frac{7}{25}})^{17}.$$

■