Variáveis Aleatórias Discretas

Pedro Henrique González

GCC 1518 – Estatística e Probabilidade – CEFET Maracanã

Probabilidade – Aplicações à Estatística

Paul L. Meyer

Introdução

Tal como nos modelos determinísticos, nos quais algumas relações funcionais desempenham importante papel (como por exemplo a linear, a quadrática, a exponencial, a trigonométrica etc.), também verificamos que, na construção de modelos não-determinísticos para fenômenos observáveis, algumas distribuições de probabilidade surgem mais freqüentemente que outras.

Um motivo para isso é que, da mesma maneira que no caso determinístico, alguns modelos matemáticos relativamente simples parecem ser capazes de descrever uma classe bastante grande de fenômenos.

A Distribuição de Poisson

Definição

Seja X uma variável aleatória discreta, tomando os seguintes valores: $0, 1, \ldots, n, \ldots$.

Se

$$P(X = k) = \frac{e^{-\alpha}\alpha^k}{k!}, \qquad k = 0, 1, \ldots, n, \ldots$$

diremos que X tem **distribuição de Poisson**, com parâmetro $\alpha > 0$.

Para verificar que esta expressão representa uma legítima distribuição de probabilidade, basta observar que

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

O Valor Esperado e a Variância da Distribuição de Poisson

Teorema

Se X tiver distribuição de Poisson, com parâmetro α, então

$$E(X) = \alpha \qquad \text{e} \qquad V(X) = \alpha.$$

Relação entre as Distribuições Binomial e de Poisson

A distribuição de Poisson representa um papel muito importante como um modelo probabilístico adequado para um grande número de fenômenos aleatórios.

Trataremos, agora, da importância dessa distribuição como uma aproximação das probabilidades binomiais.

Exemplo

Suponhamos que chamadas telefônicas cheguem a 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.

Suponhamos que pretendamos, com base nessa evidência, calcular a probabilidade de serem recebidas 0, 1, 2 etc. chamadas durante os próximos três minutos.

Exemplo

Ao considerarmos o fenômeno da chegada de chamadas, poderemos chegar à conclusão de que, a qualquer instante, uma chamada telefônica é tão provável de ocorrer como em qualquer outro instante.

Assim, a probabilidade permanece constante de momento a momento.

A dificuldade é que, mesmo em um intervalo de tempo muito pequeno, o número de pontos não é apenas infinito, mas não pode ser enumerado.

Por isso, seremos levados a uma **série de aproximações**, que descreveremos agora.

Exemplo

Para começar, poderemos considerar o intervalo de três minutos subdividido em nove intervalos de 20 segundos cada um.

Poderemos, então, tratar cada um desses nove intervalos como uma prova de Bernoulli, durante a qual observaremos uma chamada (sucesso) ou nenhuma chamada (insucesso ou falha), com

$$P(\text{sucesso}) = (1.5) \times \left(\frac{20}{60}\right) = 0.5.$$

Exemplo

Desse modo, poderemos ser tentados a afirmar que a probabilidade de duas chamadas durante o intervalo de três minutos, isto é, dois sucessos em nove tentativas, com $P(\text{sucesso}) = 0.5$, é igual a

$$\binom{9}{2}\left(\frac{1}{2}\right)^9 = \frac{9}{128}.$$

A dificuldade com esta aproximação é que estamos ignorando as possibilidades de, digamos, duas, três etc. chamadas durante um dos períodos experimentais de 20 segundos.

Se essas possibilidades fossem consideradas, o emprego da distribuição binomial não poderia ser legitimado, porque essa distribuição é aplicável somente quando existe uma **dicotomia**: uma chamada ou nenhuma chamada.

Exemplo

É para eliminar essa dificuldade que nos voltamos para a aproximação seguinte e, de fato, seremos levados a uma inteira **seqüência de aproximações**.

Uma das maneiras de estar provavelmente certo de que ao menos uma chamada será recebida na central, durante um pequeno intervalo de tempo, é fazer esse intervalo muito curto.

Assim, em vez de considerar nove intervalos de 20 segundos de duração, vamos, a seguir, considerar 18 intervalos, cada um com dez segundos de duração.

Exemplo

Agora, poderemos representar nosso experimento como 18 provas de Bernoulli, com

$$P(\text{sucesso}) = P(\text{chamada durante o subintervalo}) = (1.5)\left(\frac{10}{60}\right) = 0.25.$$

Conseqüentemente,

$$P(\text{duas chamadas durante três minutos}) = \binom{18}{2}(0.25)^2(0.75)^{16}.$$

Exemplo

Observe que, agora, estamos tratando com uma distribuição binomial diferente da anterior, isto é, temos parâmetros $n = 18$, $p = 0.25$, em lugar de $n = 9$, $p = 0.5$.

Entretanto, o valor esperado, np, é o mesmo:

$$np = 18(0.25) = 9(0.5) = 4.5.$$

Exemplo

Se continuarmos desta maneira, aumentando o número de subintervalos (isto é, n), faremos, ao mesmo tempo, decrescer a probabilidade de chegar uma chamada (isto é, p), de tal maneira que np fique **constante**.

Portanto, este exemplo nos leva a formular a seguinte pergunta: o que acontece com as probabilidades binomiais

$$\binom{n}{k} p^k (1-p)^{n-k},$$

se $n \to \infty$ e $p \to 0$, de tal maneira que np permaneça constante, digamos $np = \alpha$?

Relação entre as Distribuições Binomial e de Poisson

Teorema

Seja X uma variável aleatória distribuída binomialmente, com parâmetro p baseado em n repetições de um experimento. Isto é,

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

Admita que quando $n \to \infty$, $np = \alpha$ fique constante, ou equivalentemente, quando $n \to \infty$ e $p \to 0$, temos $np \to \alpha$. Nessas condições, teremos

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

que é a **distribuição de Poisson** com parâmetro α.

Relação entre as Distribuições Binomial e de Poisson

O teorema anterior diz, essencialmente, que poderemos obter uma aproximação das probabilidades binomiais com as probabilidades da distribuição de Poisson, toda vez que n seja grande e p seja pequeno.

Sabemos que se X tiver uma **distribuição binomial**,

$$E(X) = np,$$

enquanto que se X tiver uma **distribuição de Poisson** (com parâmetro α),

$$E(X) = \alpha.$$

Relação entre as Distribuições Binomial e de Poisson

A **distribuição binomial** é caracterizada por **dois parâmetros**, n e p, enquanto a **distribuição de Poisson** é caracterizada por **um único parâmetro**, $\alpha = np$, que representa o número esperado de sucessos por unidade de tempo (ou por unidade de espaço, em alguma outra situação).

Esse parâmetro é também conhecido com a **intensidade da distribuição**.

Já existem extensas tabulações para a distribuição de Poisson.

Exemplo

Em um cruzamento de tráfego intenso, a probabilidade p de um carro sofrer um acidente é muito pequena, digamos $p = 0.0001$.

Contudo, durante certa parte do dia, por exemplo das 16 às 18 horas, um grande número de carros passa no cruzamento (1000 carros, por exemplo).

Nessas condições, qual é a probabilidade de que dois ou mais acidentes ocorram durante este período?

Exemplo

Vamos fazer algumas hipóteses.

Devemos admitir que p seja o mesmo para todo carro.

Também devemos supor que o fato de um carro sofrer ou não um acidente não depende do que ocorra a qualquer outro carro. (Esta hipótese, obviamente, não corresponde à realidade, mas, apesar disso, a aceitaremos.)

Exemplo

Outra hipótese que estamos fazendo, não explicitamente, é que n, o número de carros que passam no cruzamento entre 16 e 18 horas, é prefixado em 1000.

Obviamente, um tratamento mais realista seria considerar o próprio n como uma variável aleatória, cujo valor depende de um mecanismo casual.

Contudo, não faremos isso e tomaremos n como pré-fixado.

Exemplo

Assim, poderemos supor que, se X for o número de acidentes entre os 1000 carros que chegam, então X terá **distribuição binomial** com $p = 0.0001$.

Deste modo, poderemos obter o valor exato da probabilidade procurada:

$$\begin{aligned} P(X=2) &= 1 - [P(X=0) + P(X=1)] \\ &= 1 - (0.9999)^{1000} - 1000(0.0001)(0.9999)^{999} \end{aligned}$$

Exemplo

Considerável dificuldade surge para o cálculo desses números.

Como n é grande e p é pequeno, poderemos aplicar o teorema e obter a seguinte aproximação:

$$P(X = k) = \frac{e^{-\alpha}\alpha^k}{k!} = \frac{e^{-0.1}(0.1)^k}{k!}.$$

Conseqüentemente,

$$P(X \geq 2) \approx 1 - e^{-0.1}(1 + 0.1) \approx 0.0045.$$

Exemplo

Suponha que um processo de fabricação produza peças de tal maneira que uma determinada proporção (constante) das peças p seja defeituosa.

Se um lote de n dessas peças for obtido, a probabilidade de encontrarmos exatamente k peças defeituosas pode ser calculada pela distribuição binomial como

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

em que X é o número de peças defeituosas no lote.

Exemplo

Se n for grande e p for pequeno (como é freqüentemente o caso), poderemos aproximar esta probabilidade por

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \approx \frac{e^{-np}(np)^k}{k!}.$$

Exemplo

Suponha, por exemplo, que um fabricante produza peças, das quais cerca de 1 em 1.000 sejam defeituosas. Isto é, $p = 0.001$.

Assim, empregando a distribuição binomial, encontraremos que, em um lote de 500 peças, a probabilidade de que nenhuma das peças seja defeituosa será $(0.999)^{500} = 0.609$.

Se aplicarmos a aproximação de Poisson, esta probabilidade poderá ser escrita como $e^{-0.5} = 0.61$.

A probabilidade de encontrar duas ou mais peças defeituosas será, de acordo com a aproximação de Poisson,

$$1 - e^{-0.5}(1 + 0.5) = 0.085.$$

Aplicações da Distribuição de Poisson

Neste curso, a distribuição de Poisson foi empregada como um recurso de aproximação de uma distribuição conhecida: a distribuição binomial.

No entanto, a distribuição de Poisson exerce por si mesma um papel extremamente importante, porque ela representa um modelo probabilístico adequado para um grande número de fenômenos observáveis.

A Distribuição Geométrica

Suponha que realizemos um experimento E e que estejamos interessados apenas na ocorrência ou não-ocorrência de algum evento A.

Admita, tal como na explicação da distribuição binomial, que realizemos E repetidamente, que as repetições sejam independentes e que, em cada repetição, $P(A) = p$ e $P(\bar{A}) = 1 - p. = q$ permaneçam os mesmos.

Suponha que repetimos o experimento até que A ocorra pela primeira vez.

Aqui nos afastamos das hipóteses que levaram à distribuição binomial, em que o número de repetições era **predeterminado**, enquanto aqui é uma **variável aleatória**.

A Distribuição Geométrica

Defina a variável aleatória X como o número de repetições necessárias para obter a primeira ocorrência de A, nele se incluindo essa última.

Assim, X toma os valores possíveis $1, 2, \ldots$.

Como $X = k$ se, e somente se, as primeiras $(k-1)$ repetições de E derem o resultado $\bar{A}$, enquanto a k-ésima repetição fornece o resultado A, temos

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

Uma variável aleatória com esta distribuição de probabilidade recebe a denominação de **distribuição geométrica**.

O Valor Esperado e a Variância da Distribuição Geométrica

Teorema

Se X tiver uma distribuição geométrica,

$$E(X) = \frac{1}{p} \qquad \text{e} \qquad V(X) = \frac{q}{p^2}.$$

Exemplo

Suponha que o custo de realização de um experimento seja R$ 1000.00.

Se o experimento falhar, ocorrerá um custo adicional de R$ 300.00, em virtude de serem necessárias algumas alterações antes que a próxima tentativa seja executada.

Se a probabilidade de sucesso em uma tentativa qualquer for 0.2, se as provas forem independentes e se os experimentos continuarem até que o primeiro resultado frutuoso seja alcançado, qual será o custo esperado do procedimento completo?

Exemplo

Se C for o custo e X for o número de provas necessárias para alcançar sucesso, teremos

$$C = 1000X + 300(X - 1) = 1300X - 300.$$

Como conseqüência,

$$E(C) = 1300E(X) - 300 = 1300\left(\frac{1}{0.2}\right) - 300 = R\$\ 6200.00$$

Exemplo

Em uma determinada localidade, a probabilidade de ocorrência de uma tormenta em algum dia durante o verão (nos meses de dezembro e janeiro) é igual a 0.1.

Admitindo independência de um dia para outro, qual é a probabilidade de ocorrência da primeira tormenta da estação de verão no dia 3 de janeiro?

Chamemos de X o número de dias (começando de primeiro de dezembro) até a primeira tormenta e façamos $P(X = 34)$, o que é igual a

$$P(X = 34) = (0.9)^{33}(0.1) = 0.003.$$

Exemplo

Se a probabilidade de que um certo ensaio dê reação positiva for igual a 0.4, qual será a probabilidade de que menos de cinco reações negativas ocorram antes da primeira positiva?

Chamando de Y o número de reações negativas antes da primeira positiva, teremos

$$P(Y = k) = (0.6)^k(0.4), \qquad k = 0, 1, 2, \ldots.$$

Logo,

$$P(Y < 5) = \sum_{k=0}^{4}(0.6)^k(0.4) = 0.92.$$

Propriedades da Distribuição Geométrica

Teorema

Suponha que X tenha uma distribuição geométrica. Então, para quaisquer inteiros positivos s e t,

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

Propriedades da Distribuição Geométrica

O teorema anterior afirma que a distribuição geométrica **não possui memória**, no seguinte sentido.

Suponha que o evento A não tenha ocorrido durante as primeiras s repetições de E.

Então, a probabilidade de que ele não ocorra durante as próximas t repetições será a mesma probabilidade de que não tivesse ocorrido durante as primeiras t repetições.

Em outras palavras, a informação de nenhum sucesso é esquecida, no que interessa aos cálculos subsequentes.

Introdução

Uma generalização óbvia da distribuição geométrica surge se propusermos a seguinte questão.

Suponha que um experimento seja continuado até que um particular evento A ocorra na r-ésima vez. Se

$$P(A) = p, \qquad P(\bar{A}) = q = 1 - p$$

em cada repetição, definiremos a variável aleatória Y como se segue: Y é o número de repetições necessárias para que A possa ocorrer exatamente r vezes.

Qual é a distribuição de probabilidade de Y?

Introdução

É evidente que, se $r = 1$, Y terá uma distribuição geométrica.

Mas $Y = k$ se, e somente se, A ocorrer na k-ésima repetição e A tiver ocorrido exatamente $(r-1)$ vezes nas $(k-1)$ repetições anteriores.

A probabilidade deste evento é meramente

$$\binom{k-1}{r-1} p^r q^{k-r},$$

pois o que acontece nas primeiras $(k-1)$ repetições é independente do que acontece na k-ésima repetição.

A Distribuição de Pascal

Portanto,

$$P(Y = k) = \binom{k-1}{r-1} p^r q^{k-r}, \qquad k = r, r+1, \ldots.$$

Vemos, facilmente, que para $r = 1$ esta expressão se reduz à expressão da distribuição geométrica.

Uma variável aleatória que tenha a distribuição de probabilidade dada por esta equação possui uma **distribuição de Pascal**.

A distribuição de Pascal é também geralmente conhecida como **distribuição binomial negativa**.

O Valor Esperado e a Variância da Distribuição de Pascal

Teorema

Se Y tiver uma distribuição de Pascal, então

$$E(Y) = \frac{r}{p}, \qquad \text{e} \qquad V(Y) = \frac{rq}{p^2}.$$

Exemplo

A probabilidade de que um experimento seja bem sucedido é 0.8.

Se o experimento for repetido até que quatro resultados bem sucedidos tenham ocorrido, qual será o número esperado de repetições necessárias?

Do que foi exposto anteriormente, teremos

$$E(\text{número de repetições}) = \frac{4}{0.8} = 5.$$

Relação entre as Distribuições Binomial e de Pascal

Suponhamos que X tenha **distribuição binomial**, com parâmetros n e p. Isto é,

$$X = \text{número de sucessos em } n \text{ provas repetidas de Bernoulli},$$

com $P(\text{sucesso}) = p$.

Suponhamos que Y tenha uma **distribuição de Pascal**, com parâmetros r e p. Isto é,

$$Y = \text{número de provas de Bernoulli necessárias para obter } r \text{ sucessos},$$

com $P(\text{sucesso}) = p$.

Relação entre as Distribuições Binomial e de Pascal

Então, valem as seguintes relações:

1. $P(Y \leq n) = P(X \geq r)$
   Se ocorrerem r ou mais sucessos nas primeiras n provas repetidas, então serão necessárias n ou menos provas para obter os primeiros r sucessos.

2. $P(Y > n) = P(X < r)$
   Se ocorrerem menos de r sucessos nas primeiras n provas, então será preciso realizar mais do que n provas para obter r sucessos.

Relação entre as Distribuições Binomial e de Pascal

As propriedades anteriores tornam possível empregar a tabulação da distribuição binomial para calcular probabilidades associadas à distribuição de Pascal.

Por exemplo, suponhamos que desejemos calcular a probabilidade de que mais de 10 repetições sejam necessárias para obter o terceiro sucesso, com $p = P(\text{sucesso}) = 0.2$.

Empregando a notação acima para X e Y, teremos

$$P(Y > 10) = P(X < 3) = \sum_{k=0}^{2} \binom{10}{k} (0.2)^k (0.8)^{10-k} = 0.678.$$

A Distribuição Hipergeométrica

Suponha que tenhamos um lote com N peças, r das quais sejam defeituosas e $(N - r)$ das quais sejam não-defeituosas.

Suponha que escolhamos, ao acaso, n peças desse lote $(n \leq N)$, sem reposição.

Seja X o número de peças defeituosas encontradas.

A Distribuição Hipergeométrica

Como $X = k$ se, e somente se, obtivermos exatamente k peças defeituosas, dentre as r defeituosas do lote, e exatamente $(n-k)$ não-defeituosas, dentre as $(N-r)$ não-defeituosas do lote, teremos

$$P(X = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}, \qquad k = 0, 1, 2, \ldots .$$

Dizemos que uma variável aleatória que tenha esta distribuição de probabilidade tem **distribuição hipergeométrica**.

Exemplo

Pequenos motores elétricos são expedidos em lotes de 50 unidades.

Antes que uma remessa seja aprovada, um inspetor escolhe cinco desses motores e os inspeciona.

Se nenhum dos motores inspecionados for defeituoso, o lote é aprovado.

Se um ou mais forem verificados defeituosos, todos os motores da remessa são inspecionados.

Suponha que existam, de fato, três motores defeituosos no lote.

Qual é a probabilidade de que a inspeção 100% seja necessária?

Exemplo

Se fizermos X igual ao número de motores defeituosos encontrados, a inspeção 100% será necessária se, e somente se, $X \geq 1$.

Logo,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{\binom{3}{0}\binom{47}{5}}{\binom{50}{5}} = 0.28.$$

O Valor Esperado e a Variância da Distribuição Hipergeométrica

Teorema

Admita que X tenha distribuição hipergeométrica. Façamos

$$p = \frac{r}{N}, \qquad q = 1 - p.$$

Nesse caso, teremos

$$E(X) = np, \qquad V(X) = npq\left(\frac{N-n}{N-1}\right) \qquad e$$

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

para N grande.

Relação entre as Distribuições Hipergeométrica e Binomial

A última propriedade do teorema afirma que, se o tamanho do lote N for suficientemente grande, a distribuição de X poderá ser aproximada pela **distribuição binomial**.

Isto é intuitivamente aceitável, porque a distribuição binomial é aplicável quando fazemos **amostragem com reposição**, pois, nesse caso, a probabilidade de obter uma peça defeituosa permanece constante, enquanto a distribuição hipergeométrica é aplicável quando fazemos **amostragem sem reposição**.

Relação entre as Distribuições Hipergeométrica e Binomial

Se o tamanho do lote for grande, não fará grande diferença se retornamos ou não ao lote uma peça determinada antes que a próxima seja escolhida.

Em geral, a aproximação da distribuição hipergeométrica pela distribuição binomial é bastante boa, se

$$\frac{n}{N} \leq 0.1.$$