Aula 5

PROBABILIDADES: TEOREMA DE BAYES, VARIÁVEL ALEATÓRIA E DISTRIBUIÇÕES DE VARIÁVEIS ALEATÓRIAS DISCRETAS

META

Estudar o comportamento e aplicação das Variáveis Aleatórias Discretas, bem como das Distribuições Binomial e Poisson.

OBJETIVOS

Ao final desta aula, o estudante deverá:

Identificar em que circunstâncias alguns experimentos podem ser observados pelo teorema de Bayes.

Saber utilizar as variáveis aleatórias discretas e calcular seus principais parâmetros: Esperança Matemática e Variância.

Aplicar as Distribuições: Binomial e Poisson a determinados tipos de experimentos de variáveis aleatórias discretas.

PRÉ-REQUISITO:

Conhecimentos da teoria de probabilidades, envolvendo: espaço amostral e principais eventos, bem como a independência de eventos. Também são importantes: Papel, Calculadora ou Computador para realização dos cálculos.

INTRODUÇÃO

Olá! Tudo bem? Vamos dar seqüência ao estudo sobre probabilidades? Para isto é fundamental tudo o que foi visto na aula anterior, novamente você vai trabalhar com espaço amostral e probabilidade de eventos.

Quando o experimento for realizado repetidamente os resultados individuais parecem ocorrer de forma acidental. Contudo, se o experimento for repetido um grande número de vezes uma configuração definida ou regularidade surgirá. É esta regularidade que torna possível construir um modelo matemático preciso, com o qual se analisará o experimento.

O desenvolvimento da teoria das probabilidades permitiu a criação de técnicas mais adequadas de amostragem e formas de relacionar as amostras e as populações de onde provieram essas amostras.

Com o desenvolvimento desta aula você vai empregar a teoria da probabilidade para tirar conclusões acerca de uma população com base em uma amostra extraída da mesma.

Com o advento da teoria das probabilidades, foi possível estabelecer as distribuições de probabilidades, consideradas à espinha dorsal da teoria estatística, pois todos os processos de inferências estatísticas são resultantes das aplicações de distribuições de probabilidades.

No caso de uma variável aleatória discreta sua distribuição de probabilidades é uma distribuição de freqüências relativas para os resultados de um espaço amostral. Estas freqüências mostram a proporção das vezes em que à referida variável tende a assumir cada um desses valores.

É importante poder descrever as características principais de uma distribuição, ou seja, caracterizar os resultados da experiência aleatória mediante alguns parâmetros. Chegamos assim, ao estudo das características associadas a uma variável aleatória, introduzindo os conceitos de esperança matemática e variância, relacionando-os com os de média e variância de uma variável estatística.

Existe uma variedade de tipos de distribuições de probabilidades na estatística. Cada qual tem seu próprio conjunto de hipótese que definem as condições sob as quais o tipo de distribuição pode ser utilizado. Nesta aula vamos trabalhar a distribuição Binomial e de Poisson.

TEOREMA DE BAYES

A ocorrência de certo acontecimento "R" pode estar condicionada a um ou mais acontecimentos (A1, A2 , ..., An), que representam as partições mutuamente exclusivos de determinado Espaço Amostral S. Dessa forma, o acontecimento "R" será igual à soma desses acontecimentos, ou seja:

$P(R) = P(A1 \cap R) + P(A2 \cap R) + ... + P(An \cap R)$

$P(R) = P(A1) * P(R / A1) + P(A2) * P(R / A2) + ... + P(An) * P(R /An)$

O teorema de Bayes, também chamado teorema das causas, relaciona-se a acontecimentos que estão associados. Imaginando a possibilidade de existir uma relação de causa e efeito entre os acontecimentos, esse teorema permitirá encontrar a probabilidade que certo evento Ai acarretou par a ocorrência de "R", isto é:

$P(Ai / R) = P(R \cap Ai) / P(R)$ Desenvolvendo e substituindo P(R) nessa equação temos:

$P(Ai / R) = (P(Ai) * P(R / Ai)) / \sum ((P(Ai) * P(R / Ai))$

Exemplo: Três máquinas C_1, C_2 e C_3 são utilizadas para fabricar arruelas. Suponha-se que cada arruela seja classificada: perfeita ou defeituosa. As máquinas produzem respectivamente 40%, 40% e 20% da produção total daquela Indústria. A percentagem de itens defeituosos da produção de C_1, C_2 e C_3 é 3%, 2% e 1%. Seleciona-se aleatoriamente um item, verificando-se ser defeituoso. Qual a probabilidade de pertencer à máquina C_2?

$P(A2 / R) = (P(A2) * P(R / A2)) / \sum ((P(Ai) * P(R / Ai))$

$P(A2 / R) = (0{,}4 * 0{,}02) / (0{,}4 * 0{,}03 + 0{,}4 * 0{,}02 + 0{,}2 * 0{,}01) = 0{,}008 / 0{,}0022 = 36{,}36\%$

VARIÁVEIS ALEATÓRIAS DISCRETAS

Uma variável aleatória é dita discreta quando ela assume somente valores inteiros num conjunto enumerável de pontos do conjunto real, isto é: é uma função que associa a cada ponto de um espaço amostral, um número real. A soma das probabilidades de todos os valores assumidos por esta variável é igual a 1.

Distribuição de Probabilidade de uma Variável Aleatória Discreta

Seja X uma Variável Aleatória Discreta definida num espaço amostral S, tal que X possa assumir valores x1, x2, ..., xn, com respectivas probabilidades p1, p2, ..., pn. A associação entre os valores assumidos por esta variável X e respectivas probabilidades de ocorrência determina a Distribuição (Função) de Probabilidade dessa variável.

A distribuição dada por P(X = xi) satisfaz as condições: **P(xi) ≥ 0 e ∑P(xi) = 1**

Exemplo: Seja X uma variável aleatória que representa o número de meninas de uma família com 4 crianças. Determine a distribuição de probabilidade de X.

S = {MMMM; MMMF; MMFM; MFMM; FMMM; MMFF; MFMF; MFFM; FFMM; FMFM; FMMF; FFFM; FFMF; FMFF; MFFF; FFFF}

Distribuição de Probabilidade de nascimento de meninas.

X = xi	*0*	*1*	*2*	*3*	*4*
pi = P(X=xi)	*1/16*	*4/16*	*6/16*	*4/16*	*1/16*

Função de Distribuição Acumulada de

Seja uma variável aleatória. Define-se a função acumulada de como:

F(X) = P(X ≥ xi), X ε R

Distribuição acumulada relativa ao nascimento de meninas do exemplo anterior:

X = xi	*0*	*1*	*2*	*3*	*4*
pi = Fxi	*1/16*	*5/16*	*11/16*	*15/16*	*1*

PARÂMETROS DE UMA VARIÁVEL ALEATÓRIA DISCRETA

1. Valor Esperado ou Esperança Matemática

Se **Xi = x1, x2, ... xn** são os possíveis valores de variável aleatória discreta e **pi** e as respectivas probabilidades da variável assumir esses valores então o **valor esperado ou esperança matemática (media),** de X, representado por $E(X)$ ou μ_x, é definido por:

E(X) = ∑ xi * pi

Considerando o nascimento de meninas do exemplo anterior o valor esperado dessa ocorrência é:

$E(X) = \mu = 0*1/16 + 1*4/16 + 2*6/16 + 3*4/16 + 4*1/16 = 32/16 = 2$

2. Variância

Seja X uma variável aleatória discreta, definimos a variância de X, denotada por V (X) ou σ_X^2, da seguinte maneira:

$V(X) = \sigma_X^2 = \sum(xi - \mu)^2 * pi.$ ou $V(X) = E(X^2) – [E(X)]^2$

Sendo: $E(X^2) = \sum xi^2 * pi$

Propriedades da Média:

$E(K) = K$ (K é constante)
$E(KX) = K * E(X)$
$E(X \pm Y) = E(X) \pm E(Y)$
$E(X \pm K) = E(X) \pm K$
$E(XY) = E(X) * E(Y)$ se X e Y são independentes
Propriedades da Variância:
$V(K) = 0$
$V(KX) = K^2 * V(X)$
$V(X \pm K) = V(X)$
$V(X + Y) = V(X) + V(Y)$ se X e Y são independentes

Considerando o nascimento de meninas do exemplo anterior a variância dessa ocorrência é:

$V(X) = \sigma_X^2 = (0 – 2)^2 * 1/16 + (1 – 2)^2 * 4/16 + (2 – 2)^2 * 6/16 + (3 – 2)^2 * 4/16 + (4 – 2)^2 * 1/16$

$V(X) = \sigma_X^2 = 4/16 + 4/16 + 0/16 + 4/16 + 4/16 = 16/16 = 1$

PRINCIPAIS DISTRIBUIÇÕES DE VARIÁVEL ALEATÓRIA DISCRETA

Distribuição Binomial

É uma distribuição de probabilidade adequada aos experimentos que apresentam apenas dois resultados (sucesso ou fracasso). Além disso, essa distribuição é apropriada para pequenas amostras (n<30), desde que o experimento investigado atenda as seguintes características:

1. O número de provas do experimento deve ser conhecido.
2. Todas as provas do experimento devem ser independentes.
3. Cada tentativa do experimento admite apenas dois resultados: sucesso ou fracasso.

Sucesso (p) – representa a probabilidade de ocorrência da característica observada.

Fracasso (q) – representa a probabilidade de que a característica observada não ocorra.

Observação: Sucesso e fracasso são eventos complementares, portanto: $p + q = 1$.

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

A variável X tem distribuição Binomial, com parâmetros n e p, e indicaremos pela notação X~B(n,p).

Ex. Em determinado Hospital 6 pacientes devem submeter-se a um tipo de operação na qual a probabilidade de êxito é de 70%. Qual a probabilidade de que 2 pacientes consigam sobreviver.

p = P(acerto) = 0,7 X : B(6, 0,7) e k = {0, 1, 2, 3, 4, 5 6}

P(X = 2) = (6! / 4! * 2!)* 0,7 * 0,7 * 0,3 * 0,3 * 0,3 * 0,3 = 15 * 0,49 * 0,0081 = 0,05953 = 5,9%

Parâmetros da Binomial

Esperança (Média) = E(X) = n.p e Var(X) = n.p.q

Distribuição de Poisson

Trata-se do caso limite da Distribuição Binomial quando o número de provas tende para infinito, enquanto a probabilidade do evento investigado tende para zero. É uma distribuição indicada para eventos independentes e raros. Essa distribuição refere-se à contagem do número de vezes que ocorre determinado sucesso por unidade de tempo (mês, dia, hora, minuto, etc.), por unidade de medida (comprimento, área, peso, etc.) ou por contagem (número de itens, de ligações, etc.).

Em muitos casos, conhece-se o número de sucessos, porém torna-se difícil e, às vezes, sem sentido determinar o número de fracassos ou o total de resultados.

Por exemplo: automóveis que passam numa esquina. Poderemos num determinado intervalo de tempo anotar o nº de carros que passaram, porém, o nº de carros que deixaram de passar pela esquina não poderá ser determinado.

Seja X o nº de sucessos no intervalo então:

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

Parâmetros de Poisson: Esperança Matemática (média) e Variância

$E(X) = np = \lambda$ e $Var(X) = np = \lambda$

<u>Ex</u>.: Num livro de 800 páginas há 800 erros de impressão. Qual a probabilidade de que uma página contenha pelo menos 3 erros?

Solução: X: nº de erros por página e $\lambda = 1$

$P(X \geq 3) = 1 – P(X < 3) = 1- \{P(X = 0) + P(X = 1) + P(X = 2)\} =$

$$1- \left\{\frac{e^{-1}.1^0}{0!} + \frac{e^{-1}.1^1}{1!} + \frac{e^{-1}.1^2}{2!}\right\} = 1 - \{0,367879 + 0,367879 + 0,183940\}$$

$= 1 – 0,919698 = 0,080302$

ATIVIDADES

1. Suponha que temos duas urnas 1 e 2, cada uma contém duas gavetas. A urna 1 contém uma moeda de ouro em uma gaveta e uma moeda de prata na outra gaveta; enquanto a urna 2 contém uma moeda de ouro em cada gaveta. Uma urna é escolhida ao acaso; a seguir uma de suas gavetas é aberta ao caso. Verifica-se que a moeda encontrada nesta gaveta é de ouro. Qual a probabilidade de que a moeda provenha da urna 2?

2. Um indivíduo pode chegar ao emprego utilizando-se apenas de um desses meios de locomoção: bicicleta, motocicleta ou carro. Sabe-se que por experiência a probabilidade de ele se utilizar do carro é de 0,6; de bicicleta 0,1 e de motocicleta 0,3. A probabilidade de chegar atrasado, no caso de se utilizar do carro, é 0,05; de motocicleta 0,08, e de bicicleta, 0,02. Certo dia ele chegou atrasado. Qual o meio de locomoção mais provável de ter sido escolhido?

3. Qual a esperança matemática e o desvio-padrão de um jogo no qual se pode ganhar R$ 25,00 com a probabilidade 0,2; R$ 10,00 com a probabilidade 0,3 e R$ 4,00 com 0,5?

4. Joga-se uma moeda três vezes. Seja X o número de caras. a) Construir a distribuição de probabilidade e a função de distribuição acumulada de X. b) Determinar E(X) e V(X).

5. Seja X uma variável aleatória que representa o número de meninas em famílias com 5 crianças. Determine a distribuição de probabilidade de X e a função de distribuição acumulada. Calcule também E(X) e V(X).

6. Um aluno conhece bem 60% da matéria dada. Num exame, com 5 perguntas sorteadas ao acaso sobre toda a matéria, que probabilidade tem de responder mais da metade das perguntas?

7. Uma companhia de seguros vendeu apólices a 5 pessoas, todas da mesma idade e com boa saúde. De acordo com as tábuas atuariais, a probabilidade de que uma pessoa daquela idade esteja viva daqui a 30 anos é de 2/3. Calcular a probabilidade de que daqui a 30 anos: a) Exatamente 2 pessoas estejam vivas; b) Todas as pessoas estejam vivas; c) Pelo menos 3 pessoas estejam vivas.

8. A probabilidade de um indivíduo apresentar cárie dentária em determinada população é de 30%. Entre seis pessoas escolhidas ao acaso, qual a probabilidade de que duas não apresentem cárie dentária?

9. A probabilidade de que um atleta ultrapasse 17,30m num único salto triplo é de 70%. O atleta dá 4 saltos. Qual a probabilidade de que pelo menos um dos saltos ultrapasse 17,30m?

10. Certo curso de treinamento aumenta a produtividade de uma população de funcionários em 80% dos casos. Se 10 funcionários quaisquer participam deste curso, encontre a probabilidade de: a) Exatamente 7 funcionários aumentarem a produtividade. b) Não mais do que 8 funcionários aumentarem a produtividade.

11.A probabilidade de um menino ser daltônico e 6%. Qual é a probabilidade de serem daltônicos todos os 4 meninos que se apresentaram, em determinado dia, para um exame oftalmológico?

12. Estima-se em 1% a probabilidade de vender apólices de seguro a pessoas que respondem a um anúncio especial. Nessa base se 800 pessoas respondem a um anúncio, qual a probabilidade de que pelo menos duas compre a referida apólice?

13. Uma pesquisa cientifica revelou que para cada 1000 pessoas entrevistadas, uma está sujeita a choque traumático quando da aplicação de penicilina. Determine a probabilidade de que, entre 700 pessoas entrevistadas ao acaso, no máximo duas sofra aquele choque nas mesmas condições.

14. A proporção de partos prematuros numa população é de 10 por 1000. Qual a probabilidade de haver no máximo 1 parto prematuro em 500 partos realizados.

15. Sendo 1% a proporção de canhotos numa população, qual a probabilidade de termos pelo menos um canhoto numa classe de 50 alunos?

16. Em uma cidade com 250.000 habitantes o número de acidentes de trânsito é de 3 por dia, em cada 100.000 habitantes. Qual a probabilidade de que ocorram, nessa cidade, em um dia mais de 2 acidentes?

CONCLUSÃO

Com esta aula, você deve ser capaz de entender o que é uma variável aleatória. Deve entender que uma variável aleatória discreta é uma função que associa a cada ponto de um espaço amostral, um número real e que a soma das probabilidades de todos os valores assumidos por esta variável é igual a um. Deve ser capaz de elaborar uma Distribuição de Probabilidade de variável aleatória discreta onde vai associar a cada valor assumido pela variável sua respectiva probabilidade de ocorrência. Com a Distribuição de probabilidade você pode calcular os parâmetros desta variável, representados pelo Valor Esperado ou Esperança Matemática e Variância, estimativas dos parâmetros calculados com base nos dados de uma população já vistos em aulas anteriores.

Entendendo o que é uma variável aleatória discreta vai utilizar a Distribuição de probabilidade Binomial que é uma distribuição de probabilidade adequada aos experimentos que apresentam apenas dois resultados (sucesso ou fracasso). Além disso, essa distribuição é apropriada para pequenas amostras e formada por eventos independentes.

Também será capaz de trabalhar com a distribuição de Poisson que está associada a experimentos cujos resultados constituem eventos independentes e que ocorram com raridade. A distribuição de Poisson é o caso limite da distribuição Binomial quando o número de provas tende para infinito, enquanto a probabilidade do evento investigado tende para zero.

No final da aula temos uma lista de exercícios para serem resolvidos em grupos de no máximo cinco pessoas ou individual. Com certeza você vai ficar muito satisfeito com os resultados do seu desempenho.

RESUMO

Nesta aula você tomou conhecimento de conceitos muito importantes na continuação dos estudos sobre probabilidades e uso de das distribuições de probabilidades. Uma variável aleatória sempre está associada à ocorrência de valores esperados, mas não certos e que esta variável pode ser discreta ou continua. Na variável aleatória discreta estes valores podem ser contados visto que seu intervalo é finito.

Utilizando dados de uma variável aleatória discreta é possível organizar os mesmos em uma distribuição de probabilidades e a partir deste conjunto calcular a Esperança Matemática e Variância. Estes parâmetros representam estimativas para a média aritmética e variância de uma população.

Dando seqüência aos temas abordados por esta aula chegamos as distribuições de freqüências e como utiliza-las em determinados experimentos. Estas distribuições foram formalizadas de acordo a ocorrência de determinadas características da variável investigada. A partir dessa associação modelos de distribuições são criados para facilitar a investigação de determinados experimentos em um valor específico ou no total do seu espaço amostral. Com a aplicação destas distribuições a maioria dos experimentos pode ser investigada de forma mais rápida, pela existência deste modelo, mas não esqueça que a base destes modelos é formada pelo espaço amostral da variável observada e os teoremas do cálculo das probabilidades, portanto os resultados de um experimento podem ser encontrados, usando o modelo estatístico apropriado ou os teoremas da probabilidade.

Entre os diversos modelos de distribuições de probabilidades de variáveis discretas você tomou conhecimento do modelo Binomial, indicado para pequenas amostras formadas por eventos independentes e o de Poisson que trabalha com grandes amostras constituídas por eventos independentes e que ocorram com raridade. Estas distribuições juntamente com a Normal, que veremos na aula seguinte, são as mais importantes para o seu curso.

AUTO-AVALIAÇÃO

Sou capaz de trabalhar com variáveis aleatórias discretas?

Sou capaz de calcular os parâmetros de uma variável aleatória discreta e entender seu significado?

Sou capaz de aplicar o modelo da distribuição de probabilidades Binomial?

Sou capaz de aplicar o modelo da distribuição de probabilidades Poisson?

PRÓXIMA AULA

Estudo das Probabilidades, envolvendo Variável Aleatória Contínua e Distribuição Normal.

REFERÊNCIAS

RODRIGUES, PEDRO CARVALHO. **Bioestatística**. Universidade Federal Fluminense.
FONSECA, JAIRO DA. **Curso de Estatística**. Editora Atlas.
OLIVEIRA, FRANCISCO ESTEVAM MARTINS DE. **Estatística e Probabilidade**. Editora Atlas.
TANAKA**. Elementos de Estatística.** Editora McGraw.Hill.
BARBETTA, PEDRO A. **Estatística aplicada às Ciências Sociais**. Editora da UFSC.
GÓES, LUIZ A. C**. Estatística I e II**. Editora Saraiva.
DÍAZ, FRANCISCA; LOPES, FRANCISCO JAVIER. **Bioestatística**. Editora Thomson.