

Estimando probabilidades

Introdução

Nas aulas anteriores estudamos o cálculo de probabilidades e aplicamos seu conceitos a vários exemplos. Assim, vimos também que nem sempre podemos calcular diretamente a probabilidade de ocorrência de um evento.

Em alguns dos problemas resolvidos, como os de jogos de azar, não foi difícil calcular a probabilidade de ocorrência de um evento utilizando nossos conhecimentos de análise combinatória. Por outro lado, em alguns problemas envolvendo pesquisas de opinião, por exemplo, o cálculo da probabilidade foi feito com base em uma amostra de uma parte da população.

Nossa aula

Nesta aula vamos discutir um pouco mais os procedimentos para o cálculo do valor da probabilidade de um evento, quanto ele não pode ser calculado diretamente ou não é conhecido previamente.

Sabemos que a maioria das indústrias possui um setor responsável pelo controle de qualidade de sua produção. Até mesmo para uma produção artesanal e/ou caseira devemos nos preocupar com o controle de qualidade do que se produz.

É claro que cada peça não aceita pelo controle de qualidade acarreta um prejuízo ao seu fabricante, seja pelo desperdício de tempo ou de material (caso essa peça não possa ser reaproveitada). Sabemos que o melhor para todos é fazer bem feito da primeira vez.

Além disso, é comum as empresas incluírem em seus preços de venda, além do custo e do lucro de uma peça, parte do prejuízo com o desperdício.

Mas como saber qual o desperdício? É preciso pelo menos *estimar* a probabilidade de fabricar uma peça rejeitada pelo controle de qualidade. Normalmente, essa é uma tarefa da equipe que trabalha nessa área.

EXEMPLO 1

Uma fábrica de camisetas tem um setor de corte da malha e outro de costura. No entanto, o fabricante está tendo prejuízo, pois várias camisetas estão sendo cortadas errado. Como não podem ser costuradas, a malha é desperdiçada. Se ele deseja controlar a qualidade de seus produtos e conhecer melhor seu problema o que deve fazer?

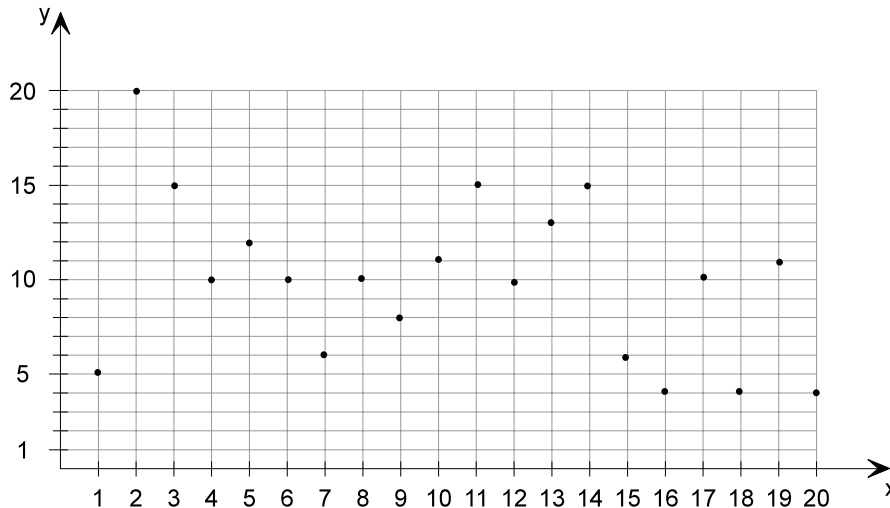
Solução 1:

Inicialmente, caso o fabricante não tenha ainda, deve criar um setor para controlar a qualidade da produção. Se sua produção é pequena a ponto de poder verificar todas as camisetas cortadas, ele não precisará trabalhar com amostras (parte da produção).

Verificando toda a produção diária, o funcionário anotará diariamente quantas camisetas são desperdiçadas. É possível que em vários dias consecutivos suas anotações tenham uma grande variação (num dia 5 e em outro 20 camisetas desperdiçadas, por exemplo).

Um procedimento muito comum seria anotar, durante vários dias consecutivos, o número de camisetas desperdiçadas e no final desse período fazer a média de todos os dias. A média aritmética do desperdício nos dará uma estimativa razoável do que ocorre diariamente.

Veja um resultado possível ao longo de **20 dias úteis** (1 mês de produção).



Pelo cálculo da média encontramos:

$$m = \frac{5 + 20 + 15 + 10 + 12 + 10 + 6}{7}$$

$$m = 10 \text{ (10 camisetas, ao dia, são desperdiçadas)}$$

Se a produção é de 100 camisetas por dia, podemos dizer que a probabilidade de encontrarmos uma camiseta defeituosa é de $10/100 = 10\%$.

Solução 2:

Outra maneira de proceder seria considerar diariamente o percentual de desperdício e depois calcularmos a média.

Solução 3:

Podemos também registrar, dia-a-dia, o número de camisetas produzidas até aquele dia e o número de camisetas defeituosas (também até aquele dia). Dividindo, dia-a-dia, o número de camisetas desperdiçadas pela produção, obteremos a porcentagem de desperdício.

dia	produção	desperdício	porcentagem
1	100	5	5%
2	200	25	12,5%
3	300	40	13,3%
4	400	50	12,5%
5	500	62	12,4%
6	600	72	12%
7	700	78	11,1%
8	800	88	11%
9	900	96	10,6%
10	1000	105	10,5%
11	1100	120	10,9%
12	1200	130	10,8%
13	1300	143	11%
14	1400	158	11,2%
15	1500	164	10,9%
16	1600	168	10,5%
17	1700	178	10,4%
18	1800	182	10,1%
19	1900	193	10,1%
20	2000	197	9,9%

Essa porcentagem é denominada *frequência relativa* de camisetas desperdiçadas.

Vemos que, nessa forma de trabalho, a partir do sétimo dia já podemos observar que os valores percentuais, ou a probabilidade de produzirmos uma camiseta defeituosa, começa a se estabilizar entre 10% e 11%.

O que nos interessa é a seqüência de freqüências relativas com que encontramos camisetas defeituosas.

É razoável supor que, se o experimento continuasse, a variação das freqüências relativas seria cada vez menor, sugerindo que, a partir de algum ponto, a freqüência se tornaria constante. Essa constante é que seria a probabilidade de fabricação de uma camiseta com defeito.

Essa forma de obtermos a probabilidade de ocorrência de um evento denomina-se *interpretação freqüentista das probabilidades*. Essa interpretação é extremamente útil para a determinação da probabilidade de eventos que podem ser testados, repetidos infinitas vezes.

EXEMPLO 2

Numa fábrica de parafusos deseja-se estimar a probabilidade de encontrar um parafuso defeituoso em sua produção de 5 mil parafusos ao dia.

Solução:

Primeiramente, vamos analisar de que forma os Exemplos 1 e 2 diferem. Observe que no caso das camisetas, além da produção diária não ser tão grande, os “erros” ocorrem no corte e, portanto, essas camisetas não são nem costuradas e, facilmente, contamos ao final do dia a quantidade de camisetas desperdiçadas. Na fábrica de parafusos, não é possível verificar diariamente todos os parafusos produzidos. Assim, o controle de qualidade precisará trabalhar com uma amostra aleatória (escolhida ao acaso) da produção diária. Observe ainda que os parafusos defeituosos são produzidos e até mesmo entregues aos clientes. Estimar a probabilidade de um cliente encontrar um parafuso defeituoso em sua encomenda é muito importante.

Os procedimentos utilizados no primeiro exemplo são os mesmos que utilizaremos agora, só que testaremos parte da produção, uma amostra. Se essa amostra (parte da produção) é escolhida aleatoriamente, isto é, ao acaso, os resultados encontrados podem ser utilizados como se tivessem sido testados todos os parafusos produzidos a cada dia.

Vamos supor que a cada 100 parafusos verificados (100 é o tamanho da amostra) contamos os defeituosos durante vários dias e descobrimos que a freqüência com que encontramos esses parafusos vai se estabilizando em $\frac{1}{20}$. Podemos então afirmar que a probabilidade de encontrarmos parafusos defeituosos é $\frac{1}{20}$ ou 5% ou podemos dizer também que a cada 20 parafusos é *provável* que encontremos 1 com defeito.

De acordo com a interpretação freqüentista das probabilidades, quando dizemos que a probabilidade é $\frac{1}{20}$, queremos dizer que a freqüência relativa segundo a qual esse evento ocorrerá numa longa seqüência de experimentações se estabilizará em $\frac{1}{20}$. É costume expressarmos esse resultado dizendo que esperamos que o evento ocorra uma vez em cada 20 experiências, ou cerca de cinco vezes em cada 100. Essa afirmação não deve ser interpretada literalmente. De forma nenhuma deve-se pensar que caso esse evento não ocorra em dezenove testagens, certamente ocorrerá na vigésima. Na vigésima testagem, a probabilidade de ocorrência desse evento continua sendo $\frac{1}{20}$.

Exercícios

Exercício 1

Lance um dado 120 vezes e observe a frequência dos resultados 1, 2, ..., 6. Os resultados obtidos estão de acordo com o que você esperava de um experimento como esse?

Exercício 2

Considere o experimento de lançarmos uma moeda. Imagine que ele foi realizado 6 000 vezes e os resultados foram anotados, constando da tabela a seguir. Complete a tabela e tire suas próprias conclusões.

Número de lançamentos	Número de caras	Frequência relativa
10	7	
20	11	
40	17	
60	24	
80	34	
100	47	
200	92	
400	204	
600	305	
800	404	
1000	492	
2000	1010	
3000	1530	
4000	2030	
5000	2517	
6000	3011	

Exercício 3

Escolha 250 números de telefone numa lista. Utilizando somente os quatro últimos algarismos de cada número - isto é, abandonando os prefixos - determine a frequência de zeros, uns, ... noes. Os dez algarismos aparecem com frequências aproximadamente iguais?