AED – 4: Recordação sobre estimação por máxima verossimilhança, aplicada a tabelas de contingência e a estatística razão de verossimilhança:

4.1 EMV

Se o sucesso ou falha de um componente em um processo de produção (*X*) pode ser modelado por uma distribuição Bernoulli (π), então *X* = 0 pode ser considerada como "falha" e *X* = 1 como "sucesso" e π como a probabilidade de sucesso. A distribuição de probabilidade para *X* é:

$$f(X) = \pi^x (1-\pi)^{1-x}$$

Para estimarmos π em um lote de produção através de uma amostra aleatória X_1, X_2, ... , X_n, podemos utilizar os resultados (x_1, x_2, ... , x_n) na função de verossimilhança, que mede a verossimilhança de diferentes valores para π. A função de verossimilhança (L) é:

$$L(\pi|x_1, x_2, \dots, x_n) = f(x_1)f(x_2) \dots f(x_n) = \prod_{i=1}^{n} f(x_i)$$

$$L(\pi|x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} \pi^{x_i} (1-\pi)^{1-x_i},$$

ou seja, a função de verossimilhança, ignorando o coeficiente da Binomial por não depender do parâmetro, será:

$$L(\pi|x_1, x_2, \dots, x_n) = \pi^{\sum x_i}(1-\pi)^{n-\sum x_i}$$

Por exemplo, se n = 10 e $\sum x_i = 3$, podemos construir o seguinte quadro para alguns valores selecionados de π e de $L(\pi|x_1, x_2, \dots, x_n)$:

π	$L(\pi\|x_1, x_2, \dots, x_n)$[1]
0,20	0,001678
0,29	0,002218
0,30	0,002224
0,31	0,002218
0,35	0,002102
0,39	0,001864
0,40	0,001792
0,41	0,001715
0,49	0,001056
0,50	0,000977

[1] Valores poderiam ser multiplicados por uma constante, no caso $C_3^{10} = 120$. Note que não há uma escala interpretativa para L.

Notamos um ponto de máxima em 0,3; ou seja, π = 0,3 é o mais provável ou mais plausível valor de π, uma vez que o mesmo maximiza a função de verossimilhança. Podemos então dizer que 0,3 é a **estimativa de máxima verossimilhança** para *π.*

A determinação deste ponto de máximo, que é o valor da **estimativa de máxima verossimilhança** para o parâmetro π, é assim obtida:

1. Aplique o logaritmo natural[2] à função de verossimilhança:

$$\ln L(\pi|x_1, x_2, \dots, x_n) = l(\pi|x_1, x_2, \dots, x_n),$$

2. Calcule a primeira derivada parcial de $l(\pi|x_1, x_2, \dots, x_n)$ em relação ao parâmetro π.

3. Iguale esta derivada à zero e resolva a equação para π, encontrando assim o estimador de máxima verossimilhança. Esta solução corresponde ao máximo de $L(\pi|x_1, x_2, \dots, x_n)$, desde que sejam atendidas certas condições de regularidade (ver, por exemplo, Mood, Graybill, Boes, 1974, pp. 276-286).

4. Calcule a derivada segunda da função encontrada e verifique se a mesma é negativa, o que corresponde a um ponto de máxima.

Para o nosso exemplo:

$$l(\pi|x_1, x_2, \dots, x_n) = \sum_{i=1}^{n} x_i \times \ln(\pi) + (n - \sum_{i=1}^{n} x_i) \times ln(1 - \pi)$$

Logo, $\frac{\partial l(\pi|x_1, x_2, \dots, x_n)}{\partial \pi} = \frac{\sum x}{\pi} - \frac{n - \sum x}{1 - \pi}$

Igualando a zero temos: $\frac{\sum x}{\pi} - \frac{n - \sum x}{1 - \pi} = 0$

O que nos dá: $\hat{\pi}$ = ...(obter o resultado e comentar sobre o estimador obtido)

[2] Teremos assim termos aditivos.

4.2 Uma forma geral de testar hipóteses: o Teste da Log Razão de Verossimilhança multiplicada por-2

A estatística razão de verossimilhança, denominada Λ, é a razão entre duas funções de verossimilhança (independente, portanto da escala). O numerador é a função de verossimilhança maximizada para o espaço de parâmetro(s) restrito pela hipótese nula. O denominador é a função de verossimilhança maximizada para o espaço de parâmetro(s) sem restrição (H_0 ou H_1) A estatística de teste[3] pode ser escrita então da seguinte forma:

$$\Lambda = \frac{máxima\ verossimilhança\ quando\ parâmetros\ satisfazem\ H_0}{máxima\ verossimilhança\ quando\ parâmetros\ satisfazem\ H_0\ \ ou\ H_1}$$

Observe que pequenos valores de Λ indicam que o valor observado é menos provável de acontecer sob a hipótese nula do que sob a hipótese alternativa.

De acordo com Wilks (1935), -2ln(Λ) pode ser aproximado por uma distribuição $\chi^2_{uG.L.}$, onde *u* corresponde à diferença entre as dimensões do espaço de parâmetros das hipóteses alternativa e nula.

No exemplo que estamos seguindo, se a hipótese nula H_0 for correspondente a π = 0,5 e a hipótese alternativa H_1 corresponder a $\pi \neq$ 0,5, a razão de verossimilhança Λ será formada, no numerador pelo (máximo) valor (possível) da função de verossimilhança sob a hipótese nula. Este máximo, para π = 0,5 é o resultado da substituição deste valor na função de verossimilhança

$$L(\pi = 0{,}5|x_1, x_2, \dots, x_n) = \ 0{,}5^{\sum x_i}(1-0{,}5)^{n-\sum x_i}$$

Sabemos que $\sum x_i = 3$ e que n = 10. Logo,

$$L(\pi = 0{,}5|x_1, x_2, \dots, x_n) = \ 0{,}5^3(1-0{,}5)^7 = 0{,}000977$$

Já o denominador da razão de verossimilhança será formado pelo máximo valor possível da função de verossimilhança sob a hipótese nula ou sob a hipótese alternativa. Em outras palavras, todos os possíveis valores de π são incluídos aqui e já sabemos que o valor máximo é 0,002224.

Logo, $\Lambda = \frac{0{,}000977}{0{,}002224} = 0{,}439299$. A estatística de teste será -2ln(Λ)= 1,6452.

Como em H_0 não há parâmetro desconhecido e em H_1 existe 1 parâmetro desconhecido, podemos afirmar que a distribuição assintótica de -2ln(Λ) segue uma $\chi^2_{1G.L.}$

Ao compararmos este valor ao valor crítico da distribuição qui-quadrado para 1 G.L. e nível de significância $\alpha = 0{,}05$ ($\chi^2_{crítico} = 3{,}84$), concluímos que não há evidência

---

[3] Note que a razão Λ independe da escala utilizada.

suficiente para rejeitar a hipótese de que π = 0,5. Ao utilizarmos *R* para encontrar o valor crítico e o "valor de p", temos:

>qchisq(0.95,1)
[1] 3.841459
>1-pchisq(1.6452,1)
[1]0.1996135

Notar que se $\Lambda = \frac{L_0}{L1}$ então:

$-2 \ln \Lambda = -2(lnL_0 - lnL_1) = 2(lnL_1 - lnL_0) = 2(l_1 - l_0)$, como às vezes a estatística é representada.

4.3 Se agora tivermos outro exemplo.
Considere a tabela abaixo, que representa nas linhas o resultado de um evento dicotômico e nas colunas 3 níveis de classificação:

	1	**2**	**3**	**Total**
0	9	3	3	**15**
1	1	2	7	**10**
Total	**10**	**5**	**10**	**25**

A hipótese nula seria, neste caso: H_0: $\pi_1 = \pi_2 = \pi_3 = \pi$

a) Qual seria a função de verossimilhança $L_{H0} = L(\pi_1, \pi_2, \pi_3) = L(\pi,\pi,\pi)$, se em H_0 temos uma proporção de sucessos (π) e, consequentemente, uma proporção de insucessos $(1 - \pi)$?

b) Para a hipótese alternativa H_1, temos a situação em que a função de verossimilhança terá seu **máximo** quando tivermos na função de verossimilhança as **proporções observadas na amostra para cada estrato (EMVs)**. Logo, teremos

$L_{H1} = L(\pi_1, \pi_2, \pi_3)$. Quais os valores estimados para π_J ?

c) Teremos então ln(Δ) = ln (L_{H0}) – ln (L_{H1}) = 10 ln(0,4) + 15 ln(0,6) – (ln(0,1)+9 ln(0,9) + 2 ln (0,4) + 3 ln (0,6) + 7 ln (0,7) + 3 ln (0,3)) = -4,1008

Logo, -2 ln(Δ) = 8,2105

d) Qual o número de graus de liberdade?

e) Qual o "valor de p" pela tabela de qui-quadrado?

f) Para o nível de significância de 0,05, qual seria a conclusão?

4.4 Exercício: Vamos agora observar a tabela de contingência formada com os seguintes dados de falhas nos anéis de pistão, retirados de um exemplo de aplicação em manutenção preventiva em engenharia de produção:

Exemplo: Falhas no anel de pistão em quatro compressores por posição no compressor:

	Norte	Centro	Sul	Total
Compr. 1	17	17	12	46
Compr. 2	11	9	13	33
Compr. 3	11	8	19	38
Compr. 4	14	7	28	49
Total	53	41	72	166

COMPRESS * POS_FALH Crosstabulation

			POS_FALH			Total
			1,00	2,00	3,00	
COMPRESS	1,00	Count	17	17	12	46
		Expected Count	14,7	11,4	20,0	46,0
		Residual	2,3	5,6	-8,0	
		Std. Residual	,6	1,7	-1,8	
		Adjusted Residual	,9	2,3	-2,8	
	2,00	Count	11	9	13	33
		Expected Count	10,5	8,2	14,3	33,0
		Residual	,5	,8	-1,3	
		Std. Residual	,1	,3	-,3	
		Adjusted Residual	,2	,4	-,5	
	3,00	Count	11	8	19	38
		Expected Count	12,1	9,4	16,5	38,0
		Residual	-1,1	-1,4	2,5	
		Std. Residual	-,3	-,5	,6	
		Adjusted Residual	-,4	-,6	,9	
	4,00	Count	14	7	28	49
		Expected Count	15,6	12,1	21,3	49,0
		Residual	-1,6	-5,1	6,7	
		Std. Residual	-,4	-1,5	1,5	
		Adjusted Residual	-,6	-2,0	2,3	
Total		Count	53	41	72	166
		Expected Count	53,0	41,0	72,0	166,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	11,722[a]	6	,068
Likelihood Ratio	12,059	6	,061
N of Valid Cases	166		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 8,15.

Como seria o cálculo da estatística -2ln(Λ), conhecido como -2 log da razão de verossimilhança, para a hipótese de independência mútua entre as posições de falha (homogeneidade das colunas)? $H_0\colon \Pi_{ij} = \Pi_{i.}\ ; i = 1,2,3,4; j = 1,2,3$

(Para ser feito e entregue, de forma escrita (impressa), até a próxima aula).

Referências Bibliográficas:

Mood, A.M, Graybill, F.A e Boes D.C. 1974. Introduction to the Theory of Statistics – 3rd Edition. International Student Edition. Tokyo: Mc-Graw-Hill Kogakusha.

Wilks, S.S. 1935. *The likelihood test of independence in contingency tables*. **Annals of Mathematical Statistics**, 6, 190-196.