Capítulo X – Teste do Qui-quadrado, χ^2

- Introdução ao qui-quadrado
- Definição geral do qui-quadrado
- Graus de liberdade e χ^2 reduzido
- Probabilidade do χ^2

166

Introdução ao Qui-Quadrado

- Consideremos um exemplo concreto. Suponhamos que realizamos 40 medidas x_1, ... x_{40} do alcance de um projéctil disparado por uma certa pistola e obtemos os resultados apresentados na tabela 10.1. Admitamos que estes resultados seguem uma distribuição normal ou Gaussiana $G_{X,\sigma}(x)$ (hipótese razoável).

Tabela 10.1 – Valores medidos de x (em cm)

731	772	771	681	722	688	653	757	733	742
739	780	709	676	760	748	672	687	766	645
678	748	689	810	805	778	764	753	709	675
698	770	754	830	725	710	738	638	787	712

- Neste tipo de experiências não conhecemos, à partida, o valor de X e a largura σ da distribuição esperada. Vamos, então, começar por estimar estas quantidades:

$$\text{Melhor estimativa para } X = \bar{x} = \frac{\sum x_i}{40} = 730.1\,\text{cm}$$

$$\text{Melhor estimativa para } \sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{39}} = 46.8\,\text{cm}$$

167

- Agora devemos perguntar: a nossa hipótese da distribuição dos 40 resultados obtidos para x corresponder a uma distribuição Gaussiana é razoável?

- Para responder a esta questão devemos calcular como esperaríamos que os 40 resultados estivessem distribuídos se correspondessem de facto a uma distribuição normal e comparar esses resultados com os experimentalmente obtidos.

- Uma dificuldade é que, como x é uma variável contínua, não podemos falar do nº esperado de medidas ser igual a um determinado valor de x. Em vez disso, devemos falar em termos do nº esperado num dado intervalo a < x < b. Vamos então dividir os valores possíveis em intervalos de valores. Os intervalos escolhidos são mostrados na tabela 10.2

Tabela 10.2 – Intervalos, k, e nº de resultados que cai em cada intervalo, O_k

Nº do intervalo k	Valores de x no intervalo		Observações O_k
1	$x < X - \sigma$	(or $x < 683.3$)	8
2	$X - \sigma < x < X$	(or $683.3 < x < 730.1$)	10
3	$X < x < X + \sigma$	(or $730.1 < x < 776.9$)	16
4	$X + \sigma < x$	(or $776.9 < x$)	6

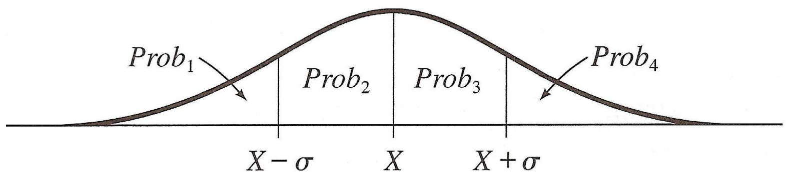
168

- Definidos os intervalos k, contamos os nºs de medidas que caem em cada intervalo, O_k.

- Em seguida, assumindo que os resultados estão normalmente distribuídos (caracterizados pelos X e σ estimados), podemos calcular o nº de medidas, E_k, que seria esperado em cada intervalo k.

- Teremos depois de decidir quão bem os nºs observados O_k se ajustam aos nºs esperados E_k

- O cálculo dos nºs esperados E_k é bastante rápido. A probabilidade de que qualquer medida caia num intervalo a < x < b é a área delimitada pela função de Gauss entre os ponto x = a e x = b. Neste exemplo, teremos as probabilidades $Prob_1$, $Prob_2$, $Prob_3$ e $Prob_4$ de uma qualquer medida cair nas áreas delimitadas pelos 4 intervalos mostrados na figura.



169

- As duas áreas iguais $Prob_2$ e $Prob_3$ correspondem à conhecida quantidade 68%. Portanto, $Prob_2 = Prob_3 = 0.34$. As outras duas probabilidades somam a quantidade remanescente 32% e, portanto, $Prob_1 = Prob_4 = 0.16$.

- Para encontrarmos os n^{os} esperados E_k temos apenas que multiplicar as quantidades E_k pelo nº total de medidas , N = 40. Os resultados são mostrados na Tabela 10.3

Tabela 10.3

Nº do intervalo, k	Probabilidade $Prob_k$	Nº esperado $E_k = N\ Prob_k$	Nº observado O_k
1	16%	6.4	8
2	34%	13.6	10
3	34%	13.6	16
4	16%	6.4	6

- Lembremos que os nºs esperados E_k, não são esperados ao fim de uma medida mas são antes valores médios depois de realizarmos o nosso "pacote" de 40 medidas muitas vezes (daí os E_k não serem nºs inteiros).

- Para avaliarmos o ajuste entre os E_k e O_k e, portanto, a justeza da nossa hipótese, determinemos os devios $(O_k – E_k)$. Se forem pequenos a hipótese estará correcta.

170

Tabela 10.4

Bin number k	Observed number O_k	Expected number $E_k = NProb_k$	Difference $O_k - E_k$
1	8	6.4	1.6
2	10	13.6	−3.6
3	16	13.6	2.4
4	6	6.4	−0.4

- Contudo, precisamos 1º de definir o que é uma diferença $(O_k – E_k)$ pequena ou grande.

- Se nos imaginarmos a repetir a nossa série de 40 medidas muitas vezes, então o nº O_k de medidas em qualquer intervalo k pode ser considerado como o resultado de uma experiência de contagens do tipo das descritas no capítulo anterior (sobre a distribuição de Poisson). As respostas muito diferentes para O_k deveriam ter como valor médio E_k e seria de esperar que fluctuassem em torno de E_k com um desvio padrão da ordem de $\sqrt{E_k}$.

- Assim, os dois nºs a serem comparados são o desvio $(O_k – E_k)$ e as fluctuações esperadas $\sqrt{E_k}$. Estas considerações levam-nos a ponderar a razão

171

$$\frac{O_k - E_k}{\sqrt{E_k}}$$

- Para alguns intervalos k, a razão deve ser positiva e para outros negativa. Para poucos intervalos k, o valor pode ser consideravelmente maior do que 1 mas para a maioria deve ser da ordem de 1 ou inferior.

- Vamos então quadrar a razão anterior para cada k e somar para todos os intervalos utilizados. Este procedimento leva-nos a definir um nº designado por qui-quadrado:

$$\chi^2 = \sum_{k=1}^{n} \frac{(O_k - E_k)^2}{E_k}$$

- Este nº constitui um indicador razoável para verificarmos o acordo entre distribuição observada e a esperada. Se $\chi^2 = 0$, o acordo é perfeito, embora essa situação seja altamente improvável. Em geral, espera-se que cada termo da soma seja da ordem de 1 e, portanto, como haverá n termos, $\chi^2 \leq n$. Essa corresponderá então à situação em que a hipótese é verificada experimentalmente. Pelo contrário, $\chi^2 >> n$ significará que a hipótese estava incorrecta.

172

- No nosso exemplo:

$$\chi^2 = \sum_{k=1}^{n} \frac{(O_k - E_k)^2}{E_k} =$$
$$= \frac{(1.6)^2}{6.4} + \frac{(-3.6)^2}{13.6} + \frac{(2.4)^2}{13.6} + \frac{(0.4)^2}{6.4} =$$
$$= 1.80$$

- Como o valor de 1.80 para χ^2 é bem menor que 4, não há razão para se duvidar que a hipótese proposta não esteja correcta

Definição geral do Qui-Quadrado

- O resultado $\chi^2 = \sum_{k=1}^{n} \frac{(O_k - E_k)^2}{E_k}$ é geral e pode ser aplicado a muitas experiência diferentes.

- A questão pertinente que surge é qual o procedimento para escolhermos o nº de intervalos. Essa escolha depende do tipo de experiência e, em particular, da quantidade x ser contínua ou discreta. É esse aspecto que vamos agora considerar.

173

Medidas de uma variável contínua

• A única distribuição contínua que estudamos foi a distribuição Gaussiana mas é claro que existem muitas mais. Por exemplo, em muitas experiências de Física Atómica e Nuclear, a distribuição esperada para a quantidade x medida (na realidade, a energia), é a distribuição Lorentziana:

$$f(x) = \frac{1}{(x-X)^2 + \gamma^2}$$

onde X e γ são certas constantes. Outro exemplo de uma distribuição contínua é a distribuição exponencial

$$f(t) = \frac{1}{\tau} e^{-t/\tau}$$

que dá a probabilidade de um átomo radioactivo (cuja vida média esperada é τ) existir durante um tempo t.

• Seja qual for a distribuição esperada f(x), a área total circunscrita pelo gráfico de f(x) em função de x é 1 e a probabilidade de obtermos uma medida entre x = a e x = b é apenas a área entre a e b.

$$\text{Prob}(a < x < b) = \int_a^b f(x)dx$$

174

• Então, se o intervalo k estiver circunscrito entre $x = a_k$ e $x = a_{k+1}$, o nº esperado de medidas no intervalo k (para um total de N medidas realizadas) é

$$\begin{aligned} E_k &= N \times \text{Prob}(a_k < x < a_{k+1}) \\ &= N\int_{a_k}^{a_{k+1}} f(x)dx \end{aligned}$$

• O nº de medidas esperadas E_k em cada intervalo não deve ser pequeno. Embora não haja limite inferior, E_k deve ser maior ou igual a 5:

$$E_k \geq 5.$$

Por outro lado, o nº de intervalos não deve ser inferior a 4, tal como no exemplo que considerámos. Desta forma, concluiremos que, para aplicarmos o teste do qui-quadrado de forma útil, o nº de medidas realizadas não deve ser inferior a 20.

Medidas de uma variável Discreta

• Suponhamos que medimos uma variável discreta ν, como por exemplo, o nº de vezes que sai o algarismo 1 quando jogamos cinco dados. Neste caso, os valores possíveis de ν são 0, 1, 2, 3, 4 e 5 e não é necessário agregarmos os valores em intervalos. Ou melhor, escolhemos 6 intervalos mas cada um deles contém apenas 1 resultado.

175

- Contudo, acontece, por vezes, ser desejável agruparmos vários resultados no mesmo intervalo. Por exemplo, se atirarmos os 5 dados 200 vezes, de acordo com o cálculo das probabilidades a distribuição dos resultados seria a apresentada na tabela 10.5.
- Como podemos ver, os números esperados quanto à obtenção de 4 ou 5 algarismos 1 é de apenas 0.6 e 0.03, respectivamente, ou seja, muito inferior às outras. Estes intervalos não estão portanto em situações adequadas à aplicação do teste do qui-quadrado. O que devemos então fazer é reagrupar os resultados ν = 3, 4 e 5 num único intervalo, ficando no total com 4 intervalos.
- A tabela 10.5 mostra também os valores esperados E_k para cada intervalo. Definidos os intervalos podemos contar as ocorrências observadas O_k em cada intervalo. Depois podemos calcular χ^2 e ver se as distribuições esperada e medida concordam.
- Nesta experiência sabemos que a distribuição esperada é certamente uma distribuição binomial $B_{5,1/6}(\nu)$, desde que os dados sejam verdadeiros. Assim, o nosso teste da distribuição é um teste a saber se os dados são verdadeiros ou estão falseados.

Tabela 10.5

Nº de vezes que sai nº 1	Ocorrências esperadas	Nº do intervalo k	Nº esperado E_k
0	80.4	1	80.4
1	80.4	2	80.4
2	32.2	3	32.2
3	6.4	4	7.0
4	0.6		
5	0.03		

176

Outras formas de Qui-quadrado

- Ao longo do curso já várias vezes apareceu a noção de χ^2 e em todos os casos o χ^2 é uma soma de quadrados de fórmula geral:

$$\chi^2 = \sum_1^n \left(\frac{valor\ observado - valor\ esperado}{desvio\ padrão} \right)^2$$

- Se o acordo for 1, χ^2 será da ordem de n; se não, será muito maior que n.
- Infelizmente, só podemos usar o teste do χ^2 quando conhecemos os valores esperados e os desvios padrão. Talvez a situação mais comum em que estes valores podem ser conhecidos com bastante precisão seja no tipo de distribuições como a que usámos como exemplo, onde os E_k são dados pela própria distribuição experimental e os desvios padrão são as raízes quadradas desses valores.
- Contudo, este teste é de aplicação muito variada. Consideremos uma situação em que medimos duas grandezas experimentais, x e y, e esperamos que haja uma determinada relação funcional entre elas, y = f(x) (como y = A + Bx). Suponhamos que temos N pares de medidas (x_i, y_i) onde os x_i têm incerteza desprezável e os y_i têm incerteza σ_i. Neste caso, os valores esperados para os y_i são os valores calculados $f(x_i)$ e podemos então testar até que ponto y se ajusta à função f(x) através de:

177

$$\chi^2 = \sum_1^N \left(\frac{y_i - f(x_i)}{\sigma_i} \right)^2$$

Graus de Liberdade e χ^2 reduzido

- Até agora comparámos o parâmetro χ^2 com o nº de intervalos em que dividimos os nossos dados. Contudo, um procedimento melhor é compará-lo antes com o nº de graus de liberdade, d.

- Em geral, o nº de graus de liberdade num cálculo estatístico é definido como o nº de dados observados *menos* o nº de parâmetros obtidos a partir desses dados e usados nos cálculos. Para os problemas usados neste capítulo os dados observados são os nºs de observações O_k nos n intervalos k = 1, … n. Portanto, o nº de dados observados é apenas n, o nº de intervalos. Então, nestes casos, o nº de graus de liberdade é

$$d = n - c$$

onde c é o nº de parâmetros que tiveram que ser calculados para determinarmos os nºs E_k esperados. Este valor **c** é frequentemente designado por *número de constrangimentos.*

178

- O nº de constrangimentos depende de cada experiência.

➢ **Lançamento de 5 dados e registo do nº de vezes em que sai o número 1**.

Nesta experiência, onde estamos a testar a hipótese dos dados serem verdadeiros, a distribuição esperada do nº de vezes que sai o número 1 (ν) é a distribuição binomial $B_{5,1/6}(\nu)$, onde n = 0, 1, 2, 3, 4 ou 5. Ambos os parâmetros desta função, o nº de dados (5) e a probabilidade de sair o número 1 (1/6), são conhecidas à partida e não têm que ser calculados a partir dos dados.

Quando calculamos o nº esperado de ocorrências de um ν particular, devemos multiplicar a probabilidade binomial pelo nº total de lançamentos, N (200 no nosso exemplo). Este parâmetro depende dos dados:

$$N = \sum_{k=1}^{n} O_k$$

Então, neste caso, c = 1 e d = n – 1 = 4 – 1 = 3, uma vez que os dados foram agrupados em 4 intervalos. Portanto, esta experiência tinha 3 graus de liberdade.

179

- **Alcance de um projéctil**

  - Os dados foram agrupados em 4 intervalos (n = 4).
  - Nessa experiência, o alcance do projéctil foi determinado N = 40 vezes.
  - Mas para determinar a distribuição $G_{X,\sigma}(x)$ calculámos X e σ a partir dos dados observados. Assim, o nº total de constrangimentos foi de 3.
  - O nº de graus de liberdade é então:

$$d = 4 - 3 = 1$$

- Este resultado explica porque não pudemos usar um nº de intervalos inferior a 4.

- Quando n é elevado, a diferença entre n e d não é importante, mas quando n é pequeno (o que acontece muitas vezes), percebe-se como é importante esta diferença.

- **Prova-se que o valor esperado para χ^2 é exactamente d**.

- Uma forma mais conveniente de avaliar o χ^2 é usar o **χ^2 reduzido**:

$$\boxed{\tilde{\chi}^2 = \frac{\chi^2}{d}}$$

$$\boxed{\text{Valor esperado de } \tilde{\chi}^2 = 1}$$

180

Probabilidades do Qui-Quadrado

- Voltemos à experiência do alcance do projéctil. Vimos que χ^2 = 1.8 e como para esta experiência d = 1, o χ^2 reduzido dá também 1.8. Será este um valor aceitável para concluirmos que a experiência é bem reproduzida por uma distribuição Gaussiana?

- O procedimento geral é o seguinte: depois de completarmos qualquer série de medidas, calculamos o χ^2 reduzido, que registaremos como $\tilde{\chi}^2_O$ uma vez que é o resulta directamente das nossas observações.

- Depois, considerando que as nossas medidas seguem uma determinada distribuição (a distribuição esperada), calculamos a probabilidade

$$\mathrm{Prob}\left(\tilde{\chi}^2 \geq \tilde{\chi}^2_O\right)$$

  de encontrarmos um valor de $\tilde{\chi}^2$ maior ou igual a $\tilde{\chi}^2_O$.

- Se esta probabilidade for elevada, o nosso valor de $\tilde{\chi}^2_O$ é perfeitamente aceitável e não há razão para rejeitarmos a distribuição esperada. Se a probabilidade for muito baixa, um valor de $\tilde{\chi}^2$ da ordem do nosso observado $\tilde{\chi}^2_O$ é muito improvável e é-o também a correspondência entre a distribuição esperada e a experimental.

181

- Uma fronteira possível são os 5% que já usámos para as correlações. Assim, uma probabilidade de

$$\mathrm{Prob}\left(\tilde{\chi}^2 \geq \tilde{\chi}_O^2\right) < 5\%$$

indicaria um desacordo significativo. Se a fronteira for 1% então uma probabilidade de

$$\mathrm{Prob}\left(\tilde{\chi}^2 \geq \tilde{\chi}_O^2\right) < 1\%$$

indicaria um desacordo altamente significativo.

- O cálculo das probabilidades $\mathrm{Prob}\left(\tilde{\chi}^2 \geq \tilde{\chi}_O^2\right)$ é complexo para o âmbito desta disciplina, mas os resultados estão tabelados e podem ser encontrados facilmente na literatura da especialidade. As percentagens da probabilidade $\mathrm{Prob}\left(\tilde{\chi}^2 \geq \tilde{\chi}_O^2\right)$ de obtermos um valor de $\tilde{\chi}^2$ maior ou igual a um valor particular $\tilde{\chi}_O^2$ para alguns graus de liberdade, d, são apresentadas da tabela 10.6. Os traços indicam probabilidades inferiores a 0.05%.

182

Tabela 10.6

	$\tilde{\chi}_o^2$												
d	0	0.25	0.5	0.75	1.0	1.25	1.5	1.75	2	3	4	5	6
1	100	62	48	39	32	26	22	19	16	8	5	3	1
2	100	78	61	47	37	29	22	17	14	5	2	0.7	0.2
3	100	86	68	52	39	29	21	15	11	3	0.7	0.2	—
5	100	94	78	59	42	28	19	12	8	1	0.1	—	—
10	100	99	89	68	44	25	13	6	3	0.1	—	—	—
15	100	100	94	73	45	23	10	4	1	—	—	—	—

Por exemplo: com 10 graus de liberdade (d = 10), a probabilidade de obter um $\tilde{\chi}^2 \geq 2$ é igual a 3%. Portanto, se tivéssemos obtido um χ^2 reduzido de 2 numa experiência de 10 graus de liberdade, concluiríamos que as nossas observações diferiam significativamente da distribuição esperada, tendo em conta um nível de significado de 5%.

183