Probabilidade e Estatística (EAD)

Tradução e adaptação:

Priscilla Priscilla Gnewuch, Márcia Helena Barbian e Maitê Mückler

Slides baseados no material desenvolvido por Mine Çetinkaya-Rundel of OpenIntro.

UFRGS UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Instituto de MATEMÁTICA E ESTATÍSTICA UFRGS

SEAD UFRGS EDUCAÇÃO A DISTÂNCIA

Capítulo 6: Inferência para dados categóricos

Slides desenvolvidos por Mine Çetinkaya-Rundel of OpenIntro.

6.1. Inferência para uma única proporção

Prática

Dois cientistas estão testando se um determinado medicamento é eficaz contra a pressão alta. O primeiro cientista quer dar o medicamento para 1000 pessoas que possuem pressão alta e, então, ver quantas delas depois de medicadas irão apresentar níveis mais baixos de pressão arterial. O segundo cientista quer medicar apenas 500 das pessoas com pressão alta e manter as outras 500 sem o medicamento, para ver quantas pessoas em ambos os grupos terão níveis baixos de pressão arterial.
Qual a melhor maneira de testar esse medicamento?

(a) Todos os 1000 recebem o medicamento.

(b) 500 recebem o medicamento e 500 não.

Prática

Dois cientistas estão testando se um determinado medicamento é eficaz contra a pressão alta. O primeiro cientista quer dar o medicamento para 1000 pessoas que possuem pressão alta e, então, ver quantas delas depois de medicadas irão apresentar níveis mais baixos de pressão arterial. O segundo cientista quer medicar apenas 500 das pessoas com pressão alta e manter as outras 500 sem o medicamento, para ver quantas pessoas em ambos os grupos terão níveis baixos de pressão arterial.
Qual a melhor maneira de testar esse medicamento?

(a) Todos os 1000 recebem o medicamento.

(b) *500 recebem o medicamento e 500 não.*

Resultados de uma pesquisa

Uma pesquisa com 670 americanos fez a mesma pergunta e abaixo está a distribuição das respostas:

Todos os 1000 recebem o medicamento	99
500 recebem o medicamento e 500 não	571
Total	670

Estimativa paramétrica e pontual

Gostaríamos de estimar a proporção de todos os americanos que possuem uma boa noção sobre desenho de experimentos, ou seja, a proporção de todos os americanos que respondem corretamente: 500 devem receber o medicamento e 500 não.

Para este caso, qual o parâmetro de interesse? E qual a estimativa pontual?

Estimativa paramétrica e pontual

Gostaríamos de estimar a proporção de todos os americanos que possuem uma boa noção sobre desenho de experimentos, ou seja, a proporção de todos os americanos que respondem corretamente: 500 devem receber o medicamento e 500 não.
Para este caso, qual o parâmetro de interesse? E qual a estimativa pontual?

- *Parâmetro de interesse:* Proporção de *todos* americanos que têm boa intuição sobre desenho de experimentos.

$$p \text{ (uma proporção populacional)}$$

Estimativa paramétrica e pontual

Gostaríamos de estimar a proporção de todos os americanos que possuem uma boa noção sobre desenho de experimentos, ou seja, a proporção de todos os americanos que respondem corretamente: 500 devem receber o medicamento e 500 não.

Para este caso, qual o parâmetro de interesse? E qual a estimativa pontual?

- *Parâmetro de interesse:* Proporção de *todos* americanos que têm boa intuição sobre desenho de experimentos.

$$p \text{ (uma proporção populacional)}$$

- *Estimativa pontual:* Proporção de americanos *amostrada* que têm boa intuição sobre desenho de experimentos.

$$\hat{p} \text{ (uma proporção da amostra)}$$

Inferência para uma proporção

Qual o percentual de todos os americanos que têm boa intuição sobre desenho de experimentos, ou seja, responderiam corretamente "500 recebem o medicamento e 500 não"?

Inferência para uma proporção

Qual o percentual de todos os americanos que têm boa intuição sobre desenho de experimentos, ou seja, responderiam corretamente "500 recebem o medicamento e 500 não"?

- Podemos responder a esta questão de pesquisa usando um intervalo de confiança, que sabemos que é sempre da forma

$$\textit{estimativa pontual} \pm ME$$

Inferência para uma proporção

Qual o percentual de todos os americanos que têm boa intuição sobre desenho de experimentos, ou seja, responderiam corretamente "500 recebem o medicamento e 500 não"?

- Podemos responder a esta questão de pesquisa usando um intervalo de confiança, que sabemos que é sempre da forma

$$estimativa\ pontual \pm ME$$

- E também sabemos que $ME = valor\ crtico \times desvio\ padrao$ da estimativa pontual.
- Desvio padrão de uma proporção amostral

$$SE_{\hat{p}} = ?$$

Inferência para uma proporção

Qual o percentual de todos os americanos que têm boa intuição sobre desenho de experimentos, ou seja, responderiam corretamente "500 recebem o medicamento e 500 não"?

- Podemos responder a esta questão de pesquisa usando um intervalo de confiança, que sabemos que é sempre da forma

$$estimativa\ pontual \pm ME$$

- E também sabemos que $ME = valor\ crtico \times desvio\ padrao$ da estimativa pontual.
- Desvio padrão de uma proporção amostral

$$SE_{\hat{p}} = ?$$

$$SE_{\hat{p}} = \sqrt{\frac{p\,(1-p)}{n}}$$

As proporções amostrais também são quase normalmente distribuídas

Teorema do limite central para proporções

As proporções amostrais serão distribuídas quase normalmente, com média igual à média da população, p e desvio padrão igual a $\sqrt{\frac{p\ (1-p)}{n}}$.

$$\hat{p} \sim N\left(mean = p, SE = \sqrt{\frac{p\ (1-p)}{n}}\right)$$

- Mas, claro, isso só é verdade sob certas condições...

Algum palpite?

As proporções amostrais também são quase normalmente distribuídas

Teorema do limite central para proporções

As proporções amostrais serão distribuídas quase normalmente, com média igual à média da população, p e desvio padrão igual a $\sqrt{\frac{p\ (1-p)}{n}}$.

$$\hat{p} \sim N\left(mean = p, SE = \sqrt{\frac{p\ (1-p)}{n}}\right)$$

- Mas, claro, isso só é verdade sob certas condições...

Algum palpite?

Observações independentes e pelo menos 10 sucessos e 10 falhas.

As proporções amostrais também são quase normalmente distribuídas

Teorema do limite central para proporções

As proporções amostrais serão distribuídas quase normalmente, com média igual à média da população, p e desvio padrão igual a $\sqrt{\frac{p\,(1-p)}{n}}$.

$$\hat{p} \sim N\left(mean = p, SE = \sqrt{\frac{p\,(1-p)}{n}}\right)$$

- Mas, claro, isso só é verdade sob certas condições...

Algum palpite?

Observações independentes e pelo menos 10 sucessos e 10 falhas.

Note: Se p é desconhecido (a maioria dos casos), usamos $\hat{p}$ no cálculo do desvio padrão.

Teorema do limite central para proporções

Suposições / condições:

1. *Independência*:
   - *Amostra aleatória*
   - *10% condição*: Se amostrar sem reposição, $n < 10$ % da população.
2. *Normalidade*: Pelo menos 10 sucessos e 10 falhas.

De volta ao desenho de experimentos...

A pesquisa com americanos descobriu que 571 das 670 (85%) pessoas responderam corretamente à questão do desenho de experimentos.

Qual a estimativa (usando um intervalo de confiança de 95%) da proporção de todos os americanos que têm boa intuição sobre o desenho de experimentos?

De volta ao desenho de experimentos...

A pesquisa com americanos descobriu que 571 das 670 (85%) pessoas responderam corretamente à questão do desenho de experimentos.

Qual a estimativa (usando um intervalo de confiança de 95%) da proporção de todos os americanos que têm boa intuição sobre o desenho de experimentos?

Dados: $n = 670, \hat{p} = 0.85$. Primeiro, vamos verificar se as condições estão atendidas:

De volta ao desenho de experimentos...

A pesquisa com americanos descobriu que 571 das 670 (85%) pessoas responderam corretamente à questão do desenho de experimentos.

Qual a estimativa (usando um intervalo de confiança de 95%) da proporção de todos os americanos que têm boa intuição sobre o desenho de experimentos?

Dados: $n = 670, \hat{p} = 0.85$. Primeiro, vamos verificar se as condições estão atendidas:

1. *Independência*: A amostra é aleatória e 670 $<$ 10% de todos os americanos, portanto, podemos supor que a resposta de um entrevistado é independente de outra resposta.

De volta ao desenho de experimentos...

A pesquisa com americanos descobriu que 571 das 670 (85%) pessoas responderam corretamente à questão do desenho de experimentos.

Qual a estimativa (usando um intervalo de confiança de 95%) da proporção de todos os americanos que têm boa intuição sobre o desenho de experimentos?

Dados: $n = 670, \hat{p} = 0.85$. Primeiro, vamos verificar se as condições estão atendidas:

1. *Independência*: A amostra é aleatória e 670 $<$ 10% de todos os americanos, portanto, podemos supor que a resposta de um entrevistado é independente de outra resposta.
2. *Sucesso-falha*: 571 pessoas responderam corretamente (sucessos) e 99 responderam incorretamente (falhas), ambas as quantidades são maiores que 10.

Prática

Temos que $n = 670, \hat{p} = 0.85$ e, além disso, acabamos de aprender que o desvio padrão de uma proporção amostral é $SE = \sqrt{\frac{p(1-p)}{n}}$. Qual das alternativas abaixo é o cálculo correto do intervalo de confiança de 95% para este caso?

(a) $0.85 \pm 1.96 \times \sqrt{\frac{0.85 \times 0.15}{670}}$

(b) $0.85 \pm 1.65 \times \sqrt{\frac{0.85 \times 0.15}{670}}$

(c) $0.85 \pm 1.96 \times \frac{0.85 \times 0.15}{\sqrt{670}}$

(d) $571 \pm 1.96 \times \sqrt{\frac{571 \times 99}{670}}$

Prática

Temos que $n = 670, \hat{p} = 0.85$ e, além disso, acabamos de aprender que o desvio padrão de uma proporção amostral é $SE = \sqrt{\frac{p(1-p)}{n}}$. Qual das alternativas abaixo é o cálculo correto do intervalo de confiança de 95% para este caso?

(a) $0.85 \pm 1.96 \times \sqrt{\frac{0.85 \times 0.15}{670}} \rightarrow (0.82, 0.88)$

(b) $0.85 \pm 1.65 \times \sqrt{\frac{0.85 \times 0.15}{670}}$

(c) $0.85 \pm 1.96 \times \frac{0.85 \times 0.15}{\sqrt{670}}$

(d) $571 \pm 1.96 \times \sqrt{\frac{571 \times 99}{670}}$

Interpretação do IC

Com base nesse intervalo de confiança, podemos dizer que aparentemente mais de 80% dos americanos têm uma boa intuição sobre desenho de experimentos?

(0.82, 0.88)

(a) Sim

(b) Não

(c) Não sabe

Interpretação do IC

Com base nesse intervalo de confiança, podemos dizer que aparentemente mais de 80% dos americanos têm uma boa intuição sobre desenho de experimentos?

(0.82, 0.88)

(a) *Sim*

(b) Não

(c) Não sabe

Escolhendo um tamanho de amostra

Quantas pessoas você deve amostrar para reduzir a margem de erro de um intervalo de confiança de 95% para 1%?

Escolhendo um tamanho de amostra

Quantas pessoas você deve amostrar para reduzir a margem de erro de um intervalo de confiança de 95% para 1%?

$$ME = z^{\star} \times SE$$

Escolhendo um tamanho de amostra

Quantas pessoas você deve amostrar para reduzir a margem de erro de um intervalo de confiança de 95% para 1%?

$$ME = z^{\star} \times SE$$

$$0.01 \quad \geq \quad 1.96 \times \sqrt{\frac{0.85 \times 0.15}{n}} \rightarrow \textit{Use estimativa para } \hat{p} \textit{ de estudo anterior}$$

Escolhendo um tamanho de amostra

Quantas pessoas você deve amostrar para reduzir a margem de erro de um intervalo de confiança de 95% para 1%?

$$ME = z^{\star} \times SE$$

$$\begin{aligned} 0.01 &\geq 1.96 \times \sqrt{\frac{0.85 \times 0.15}{n}} \rightarrow \textit{Use estimativa para } \hat{p} \textit{ de estudo anterior} \\ 0.01^2 &\geq 1.96^2 \times \frac{0.85 \times 0.15}{n} \end{aligned}$$

Escolhendo um tamanho de amostra

Quantas pessoas você deve amostrar para reduzir a margem de erro de um intervalo de confiança de 95% para 1%?

$$ME = z^{\star} \times SE$$

$$\begin{aligned} 0.01 &\geq 1.96 \times \sqrt{\frac{0.85 \times 0.15}{n}} \rightarrow \textit{Use estimativa para } \hat{p} \textit{ de estudo anterior} \\ 0.01^2 &\geq 1.96^2 \times \frac{0.85 \times 0.15}{n} \\ n &\geq \frac{1.96^2 \times 0.85 \times 0.15}{0.01^2} \end{aligned}$$

Escolhendo um tamanho de amostra

Quantas pessoas você deve amostrar para reduzir a margem de erro de um intervalo de confiança de 95% para 1%?

$$ME = z^{\star} \times SE$$

$$\begin{aligned}
0.01 &\geq 1.96 \times \sqrt{\frac{0.85 \times 0.15}{n}} \rightarrow \textit{Use estimativa para } \hat{p} \textit{ de estudo anterior} \\
0.01^2 &\geq 1.96^2 \times \frac{0.85 \times 0.15}{n} \\
n &\geq \frac{1.96^2 \times 0.85 \times 0.15}{0.01^2} \\
n &\geq 4898.04
\end{aligned}$$

Escolhendo um tamanho de amostra

Quantas pessoas você deve amostrar para reduzir a margem de erro de um intervalo de confiança de 95% para 1%?

$$ME = z^{\star} \times SE$$

$$\begin{aligned}
0.01 &\geq 1.96 \times \sqrt{\frac{0.85 \times 0.15}{n}} \rightarrow \textit{Use estimativa para } \hat{p} \textit{ de estudo anterior} \\
0.01^2 &\geq 1.96^2 \times \frac{0.85 \times 0.15}{n} \\
n &\geq \frac{1.96^2 \times 0.85 \times 0.15}{0.01^2} \\
n &\geq 4898.04 \rightarrow \textit{n deve ser pelo menos 4,899}
\end{aligned}$$

E se não houver um estudo anterior?

Use $\hat{p} = 0.5$

Por quê?

E se não houver um estudo anterior?

Use $\hat{p} = 0.5$

Por quê?

- se você não sabe nada sobre uma proporção, 50-50 é um bom palpite!

E se não houver um estudo anterior?

Use $\hat{p} = 0.5$

Por quê?

- se você não sabe nada sobre uma proporção, 50-50 é um bom palpite!
- $\hat{p} = 0.5$ fornece uma estimativa mais conservadora - maior tamanho de amostra possível.

IC vs. TH para proporções

- Condição de sucesso-falha:
  - IC: Pelo menos 10 sucessos e falhas *observados*.
  - TH: Pelo menos 10 sucessos e falhas *esperados*, calculados usando a hipótese nula.
- Desvio padrão:
  - IC: calcular usando a proporção amostral observada: $SE = \sqrt{\frac{p(1-p)}{n}}$
  - TH: calcular usando a hipótese nula: $SE = \sqrt{\frac{p_0(1-p_0)}{n}}$

Prática

A pesquisa com americanos descobriu que 571 das 670 (85%) pessoas responderam corretamente à questão do desenho de experimentos.

Qual a estimativa (usando um intervalo de confiança de 95%) da proporção de todos os americanos que têm boa intuição sobre o desenho de experimentos?

Prática

A pesquisa com americanos descobriu que 571 das 670 (85%) pessoas responderam corretamente à questão do desenho de experimentos.

Qual a estimativa (usando um intervalo de confiança de 95%) da proporção de todos os americanos que têm boa intuição sobre o desenho de experimentos?

$$H_0 : p = 0.80 \qquad H_A : p > 0.80$$

Prática

A pesquisa com americanos descobriu que 571 das 670 (85%) pessoas responderam corretamente à questão do desenho de experimentos.

Qual a estimativa (usando um intervalo de confiança de 95%) da proporção de todos os americanos que têm boa intuição sobre o desenho de experimentos?

$$H_0 : p = 0.80 \qquad H_A : p > 0.80$$

$$SE = \sqrt{\frac{0.80 \times 0.20}{670}} = 0.0154$$

Prática

A pesquisa com americanos descobriu que 571 das 670 (85%) pessoas responderam corretamente à questão do desenho de experimentos.

Qual a estimativa (usando um intervalo de confiança de 95%) da proporção de todos os americanos que têm boa intuição sobre o desenho de experimentos?

$$H_0 : p = 0.80 \qquad H_A : p > 0.80$$

$$\begin{aligned} SE &= \sqrt{\frac{0.80 \times 0.20}{670}} = 0.0154 \\ Z &= \frac{0.85 - 0.80}{0.0154} = 3.25 \end{aligned}$$

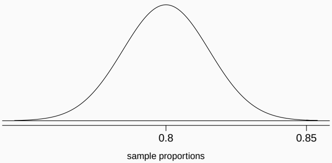
Prática

A pesquisa com americanos descobriu que 571 das 670 (85%) pessoas responderam corretamente à questão do desenho de experimentos.

Qual a estimativa (usando um intervalo de confiança de 95%) da proporção de todos os americanos que têm boa intuição sobre o desenho de experimentos?

$$H_0 : p = 0.80 \qquad H_A : p > 0.80$$

$$\begin{aligned} SE &= \sqrt{\frac{0.80 \times 0.20}{670}} = 0.0154 \\ Z &= \frac{0.85 - 0.80}{0.0154} = 3.25 \\ valor - p &= 1 - 0.9994 = 0.0006 \end{aligned}$$



Prática

Como o o-valor é baixo, rejeitamos H_0. Os dados fornecem evidências convincentes de que mais de 80% dos americanos têm uma boa intuição sobre desenho de experimentos.

Prática

Em uma pesquisa do Gallup de 2006, 11% de 1.001 americanos afirmaram que possuem resistência para celebrar o Halloween por motivos religiosos. Ao nível de confiança de 95%, a margem de erro dessa pesquisa é de $\pm 3\%$.

Uma notícia, ao descobrir o resultado dessa pesquisa, publicou: "Mais de 10% de todos os americanos possuem resistência para celebrar o Halloween por motivos religiosos".

Ao nível de confiança de 95%, a declaração feita nessa notícia é justificada?

(a) Sim

(b) Não

(c) Nao sabe

Prática

Em uma pesquisa do Gallup de 2006, 11% de 1.001 americanos afirmaram que possuem resistência para celebrar o Halloween por motivos religiosos. Ao nível de confiança de 95%, a margem de erro dessa pesquisa é de $\pm 3\%$.

Uma notícia, ao descobrir o resultado dessa pesquisa, publicou: "Mais de 10% de todos os americanos possuem resistência para celebrar o Halloween por motivos religiosos".

Ao nível de confiança de 95%, a declaração feita nessa notícia é justificada?

(a) Sim

(b) *Não*

(c) Nao sabe

Recapitulando - inferência para uma proporção

- Parâmetro da população: p, estimativa pontual: $\hat{p}$

Recapitulando - inferência para uma proporção

- Parâmetro da população: p, estimativa pontual: $\hat{p}$
- Condições:
  - independência
    - amostra aleatória e condição 10%
  - pelo menos 10 sucessos e fracassos
    - se não $\rightarrow$ aleatorizar

Recapitulando - inferência para uma proporção

- Parâmetro da população: p, estimativa pontual: $\hat{p}$
- Condições:
  - independência
    - amostra aleatória e condição 10%
  - pelo menos 10 sucessos e fracassos
    - se não $\rightarrow$ aleatorizar
- Desvio Padrão: $SE = \sqrt{\frac{p(1-p)}{n}}$
  - para IC: usar $\hat{p}$
  - para TH: usar p_0

6.2. Diferença entre proporções

Derretimento de calota de gelo

Cientistas preveem que o aquecimento global pode ter grandes consequências nas regiões polares nos próximos 100 anos. Um dos possíveis efeitos é que a calota de gelo do polo norte pode derreter completamente. Se isso acontecesse realmente, te incomodaria muito, um pouco, quase nada ou nada?

(a) Muito

(b) Um pouco

(c) Quase nada

(d) Nada

Resultados de uma pesquisa

Uma pesquisa americana faz a mesma pergunta e abaixo estão as distribuições de respostas. Além disso, essa mesma pergunta foi feita para um grupo de estudantes de estatística introdutória da Universidade de Duke:

	Pesquisa EUA	Duke
Muito	454	69
Um pouco	124	30
Quase nada	52	4
Nada	50	2
Total	680	105

Estimativa paramétrica e pontual

- *Parâmetro de interesse:* Diferença entre as proporções de *todos* os alunos da Duke e *todos* os americanos que ficariam muito incomodados com a calota de gelo do polo norte derretendo completamente.

$$p_{Duke} - p_{EUA}$$

Estimativa paramétrica e pontual

- *Parâmetro de interesse:* Diferença entre as proporções de *todos* os alunos da Duke e *todos* os americanos que ficariam muito incomodados com a calota de gelo do polo norte derretendo completamente.

$$p_{Duke} - p_{EUA}$$

- *Estimação pontual:* Diferença entre as proporções da *amostra* de estudantes da Duke e da *amostra* de americanos que ficariam muito incomodados com a calota de gelo do polo norte derretendo completamente.

$$\hat{p}_{Duke} - \hat{p}_{EUA}$$

Inferência para comparar proporções

- Os detalhes são os mesmos de antes...

Inferência para comparar proporções

- Os detalhes são os mesmos de antes...
- IC: *estimativa pontual* $\pm$ *margem de erro*

Inferência para comparar proporções

- Os detalhes são os mesmos de antes...
- IC: $estimativa\ pontual \pm margem\ de\ erro$
- TH: Usar $Z = \frac{estimativa\ pontual - valor\ nulo}{SE}$ para encontrar o p-valor apropriado.

Inferência para comparar proporções

- Os detalhes são os mesmos de antes...
- IC: $\textit{estimativa pontual} \pm \textit{margem de erro}$
- TH: Usar $Z = \frac{\textit{estimativa pontual} - \textit{valor nulo}}{SE}$ para encontrar o p-valor apropriado.
- Nós só precisamos do desvio padrão apropriado da estimativa pontual ($SE_{\hat{p}_{Duke} - \hat{p}_{EUA}}$), que é o único novo conceito apresentado aqui.

Inferência para comparar proporções

- Os detalhes são os mesmos de antes...
- IC: *estimativa pontual* $\pm$ *margem de erro*
- TH: Usar $Z = \frac{estimativa\ pontual - valor\ nulo}{SE}$ para encontrar o p-valor apropriado.
- Nós só precisamos do desvio padrão apropriado da estimativa pontual ($SE_{\hat{p}_{Duke} - \hat{p}_{EUA}}$), que é o único novo conceito apresentado aqui.

Desvio padrão da diferença entre duas proporções da amostra

$$SE_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

Condições para IC para diferença de proporções

1. *Independência dentro dos grupos:*
    - O grupo dos EUA é amostrado aleatoriamente e estamos assumindo que o grupo Duke representa uma amostra aleatória também.

Condições para IC para diferença de proporções

1. *Independência dentro dos grupos:*
   - O grupo dos EUA é amostrado aleatoriamente e estamos assumindo que o grupo Duke representa uma amostra aleatória também.
   - $n_{Duke} < 10\%$ de todos os alunos da Duke e $680 < 10\ \%$ de todos os americanos.

Condições para IC para diferença de proporções

1. *Independência dentro dos grupos:*
   - O grupo dos EUA é amostrado aleatoriamente e estamos assumindo que o grupo Duke representa uma amostra aleatória também.
   - $n_{Duke} < 10\%$ de todos os alunos da Duke e $680 < 10\ \%$ de todos os americanos.

   Podemos supor que as respostas dos alunos da Duke na amostra são independentes umas das outras, e as respostas dos americanos na amostra são independentes umas das outras também.

Condições para IC para diferença de proporções

1. *Independência dentro dos grupos:*
    - O grupo dos EUA é amostrado aleatoriamente e estamos assumindo que o grupo Duke representa uma amostra aleatória também.
    - $n_{Duke} < 10\%$ de todos os alunos da Duke e $680 < 10$ % de todos os americanos.

    Podemos supor que as respostas dos alunos da Duke na amostra são independentes umas das outras, e as respostas dos americanos na amostra são independentes umas das outras também.
2. *Independência entre grupos:* Os alunos amostrados da Duke e os americanos na pesquisa EUA são independentes uns dos outros.

Condições para IC para diferença de proporções

1. *Independência dentro dos grupos:*
   - O grupo dos EUA é amostrado aleatoriamente e estamos assumindo que o grupo Duke representa uma amostra aleatória também.
   - $n_{Duke} < 10\%$ de todos os alunos da Duke e $680 < 10$ % de todos os americanos.

   Podemos supor que as respostas dos alunos da Duke na amostra são independentes umas das outras, e as respostas dos americanos na amostra são independentes umas das outras também.
2. *Independência entre grupos:* Os alunos amostrados da Duke e os americanos na pesquisa EUA são independentes uns dos outros.
3. *Sucesso-falha:* Pelo menos 10 sucessos observados e 10 falhas observadas nos dois grupos.

Prática

Construa um intervalo de confiança de 95% para a diferença entre as proporções de alunos da Duke e de americanos que seriam muito incomodados com o derretimento da calota de gelo do polo norte $(p_{Duke} - p_{EUA})$.

Dados	Duke	EUA
Muito	69	454
Nada	36	226
Total	105	680

Prática

Construa um intervalo de confiança de 95% para a diferença entre as proporções de alunos da Duke e de americanos que seriam muito incomodados com o derretimento da calota de gelo do polo norte $(p_{Duke} - p_{EUA})$.

Dados	Duke	EUA
Muito	69	454
Nada	36	226
Total	105	680
$\hat{p}$	0.657	0.668

Prática

$$(\hat{p}_{Duke} - \hat{p}_{EUA}) \pm z^{\star} \times \sqrt{\frac{\hat{p}_{Duke}(1-\hat{p}_{Duke})}{n_{Duke}} + \frac{\hat{p}_{EUA}(1-\hat{p}_{EUA})}{n_{EUA}}}$$

Prática

$$
\begin{aligned}
& (\hat{p}_{Duke} - \hat{p}_{EUA}) \pm z^{\star} \times \sqrt{\frac{\hat{p}_{Duke}(1 - \hat{p}_{Duke})}{n_{Duke}} + \frac{\hat{p}_{EUA}(1 - \hat{p}_{EUA})}{n_{EUA}}} \\
= \quad & (0.657 - 0.668)
\end{aligned}
$$

Prática

$$
\begin{aligned}
& (\hat{p}_{Duke} - \hat{p}_{EUA}) \pm z^{\star} \times \sqrt{\frac{\hat{p}_{Duke}(1 - \hat{p}_{Duke})}{n_{Duke}} + \frac{\hat{p}_{EUA}(1 - \hat{p}_{EUA})}{n_{EUA}}} \\
= \quad & (0.657 - 0.668) \pm 1.96
\end{aligned}
$$

Prática

$$
\begin{aligned}
& (\hat{p}_{Duke} - \hat{p}_{EUA}) \pm z^{\star} \times \sqrt{\frac{\hat{p}_{Duke}(1-\hat{p}_{Duke})}{n_{Duke}} + \frac{\hat{p}_{EUA}(1-\hat{p}_{EUA})}{n_{EUA}}} \\
= \quad & (0.657 - 0.668) \pm 1.96 \times \sqrt{\frac{0.657 \times 0.343}{105} + \frac{0.668 \times 0.332}{680}}
\end{aligned}
$$

Prática

$$
\begin{aligned}
& (\hat{p}_{Duke} - \hat{p}_{EUA}) \pm z^{\star} \times \sqrt{\frac{\hat{p}_{Duke}(1-\hat{p}_{Duke})}{n_{Duke}} + \frac{\hat{p}_{EUA}(1-\hat{p}_{EUA})}{n_{EUA}}} \\
= \quad & (0.657 - 0.668) \pm 1.96 \times \sqrt{\frac{0.657 \times 0.343}{105} + \frac{0.668 \times 0.332}{680}} \\
= \quad & -0.011 \pm
\end{aligned}
$$

Prática

$$
\begin{aligned}
& (\hat{p}_{Duke} - \hat{p}_{EUA}) \pm z^\star \times \sqrt{\frac{\hat{p}_{Duke}(1 - \hat{p}_{Duke})}{n_{Duke}} + \frac{\hat{p}_{EUA}(1 - \hat{p}_{EUA})}{n_{EUA}}} \\
= \quad & (0.657 - 0.668) \pm 1.96 \times \sqrt{\frac{0.657 \times 0.343}{105} + \frac{0.668 \times 0.332}{680}} \\
= \quad & -0.011 \pm 1.96 \times 0.0497
\end{aligned}
$$

Prática

$$
\begin{aligned}
& (\hat{p}_{Duke} - \hat{p}_{EUA}) \pm z^{\star} \times \sqrt{\frac{\hat{p}_{Duke}(1-\hat{p}_{Duke})}{n_{Duke}} + \frac{\hat{p}_{EUA}(1-\hat{p}_{EUA})}{n_{EUA}}} \\
= \quad & (0.657 - 0.668) \pm 1.96 \times \sqrt{\frac{0.657 \times 0.343}{105} + \frac{0.668 \times 0.332}{680}} \\
= \quad & -0.011 \pm 1.96 \times 0.0497 \\
= \quad & -0.011 \pm 0.097
\end{aligned}
$$

Prática

$$
\begin{aligned}
& (\hat{p}_{Duke} - \hat{p}_{EUA}) \pm z^{\star} \times \sqrt{\frac{\hat{p}_{Duke}(1-\hat{p}_{Duke})}{n_{Duke}} + \frac{\hat{p}_{EUA}(1-\hat{p}_{EUA})}{n_{EUA}}} \\
= \quad & (0.657 - 0.668) \pm 1.96 \times \sqrt{\frac{0.657 \times 0.343}{105} + \frac{0.668 \times 0.332}{680}} \\
= \quad & -0.011 \pm 1.96 \times 0.0497 \\
= \quad & -0.011 \pm 0.097 \\
= \quad & (-0.108, 0.086)
\end{aligned}
$$

Prática

Qual dos seguintes conjuntos de hipóteses é o correto para testar se a proporção de pessoas que ficariam muito incomodados pelo derretimento da calota de gelo do polo norte entre os alunos da Duke difere da proporção de todos os americanos que também ficariam muito incomodados?

(a) $H_0 : p_{Duke} = p_{EUA}$
$H_A : p_{Duke} \neq p_{US}$

(b) $H_0 : \hat{p}_{Duke} = \hat{p}_{EUA}$
$H_A : \hat{p}_{Duke} \neq \hat{p}_{EUA}$

(c) $H_0 : p_{Duke} - p_{EUA} = 0$
$H_A : p_{Duke} - p_{EUA} \neq 0$

(d) $H_0 : p_{Duke} = p_{EUA}$
$H_A : p_{Duke} < p_{EUA}$

Prática

Qual dos seguintes conjuntos de hipóteses é o correto para testar se a proporção de pessoas que ficariam muito incomodados pelo derretimento da calota de gelo do polo norte entre os alunos da Duke difere da proporção de todos os americanos que também ficariam muito incomodados?

(a) $H_0 : p_{Duke} = p_{EUA}$

$H_A : p_{Duke} \neq p_{US}$

(b) $H_0 : \hat{p}_{Duke} = \hat{p}_{EUA}$

$H_A : \hat{p}_{Duke} \neq \hat{p}_{EUA}$

(c) $H_0 : p_{Duke} - p_{EUA} = 0$

$H_A : p_{Duke} - p_{EUA} \neq 0$

(d) $H_0 : p_{Duke} = p_{EUA}$

$H_A : p_{Duke} < p_{EUA}$

Ambos (a) e (c) estão corretas.

Flashback para trabalhar com uma proporção

- Ao construir um intervalo de confiança para uma proporção da população, verificamos se o número de sucessos e falhas *observado* é pelo menos 10.

$$n\hat{p} \geq 10 \qquad\qquad n(1 - \hat{p}) \geq 10$$

Flashback para trabalhar com uma proporção

- Ao construir um intervalo de confiança para uma proporção da população, verificamos se o número de sucessos e falhas *observado* é pelo menos 10.

$$n\hat{p} \geq 10 \qquad\qquad n(1-\hat{p}) \geq 10$$

- Ao conduzir um teste de hipótese para uma proporção da população, verificamos se o número de sucessos e falhas *esperado* é de pelo menos 10.

$$np_0 \geq 10 \qquad\qquad n(1-p_0) \geq 10$$

Estimativa agrupada de uma proporção

- No caso de comparar duas proporções onde $H_0 : p_1 = p_2$, não há um valor nulo que possamos usar para calcular o número de sucessos e falhas *esperado* em cada amostra.

Estimativa agrupada de uma proporção

- No caso de comparar duas proporções onde $H_0 : p_1 = p_2$, não há um valor nulo que possamos usar para calcular o número de sucessos e falhas *esperado* em cada amostra.
- Portanto, precisamos primeiro encontrar uma proporção comum (*agrupado*) para os dois grupos e usá-la em nossa análise.

Estimativa agrupada de uma proporção

- No caso de comparar duas proporções onde $H_0 : p_1 = p_2$, não há um valor nulo que possamos usar para calcular o número de sucessos e falhas *esperado* em cada amostra.
- Portanto, precisamos primeiro encontrar uma proporção comum (*agrupado*) para os dois grupos e usá-la em nossa análise.
- Isso significa encontrar a proporção de sucessos totais entre o número total de observações.

Estimativa agrupada de uma proporção

$$\hat{p} = \frac{\#\ de\ sucessos_1 + \#\ de\ sucessos_2}{n_1 + n_2}$$

Prática

Calcular a estimativa proporção combinada de estudantes da Duke e americanos que se incomodariam muito com o derretimento da calota de gelo do polo norte. De qual proporção amostrada ($\hat{p}_{Duke}$ ou $\hat{p}_{EUA}$) a estimativa agrupada está mais próxima? Por quê?

Data	Duke	EUA
Muito	69	454
Nada	36	226
Total	105	680
$\hat{p}$	0.657	0.668

Prática

Calcular a estimativa proporção combinada de estudantes da Duke e americanos que se incomodariam muito com o derretimento da calota de gelo do polo norte. De qual proporção amostrada ($\hat{p}_{Duke}$ ou $\hat{p}_{EUA}$) a estimativa agrupada está mais próxima? Por quê?

Data	Duke	EUA
Muito	69	454
Nada	36	226
Total	105	680
$\hat{p}$	0.657	0.668

$$\hat{p} = \frac{\# \; de \; sucessos_1 + \# \; de \; sucessos_2}{n_1 + n_2}$$

Prática

Calcular a estimativa proporção combinada de estudantes da Duke e americanos que se incomodariam muito com o derretimento da calota de gelo do polo norte. De qual proporção amostrada ($\hat{p}_{Duke}$ ou $\hat{p}_{EUA}$) a estimativa agrupada está mais próxima? Por quê?

Data	Duke	EUA
Muito	69	454
Nada	36	226
Total	105	680
$\hat{p}$	0.657	0.668

$$\hat{p} = \frac{\#\ de\ sucessos_1 + \#\ de\ sucessos_2}{n_1 + n_2} = \frac{69 + 454}{105 + 680}$$

Prática

Calcular a estimativa proporção combinada de estudantes da Duke e americanos que se incomodariam muito com o derretimento da calota de gelo do polo norte. De qual proporção amostrada ($\hat{p}_{Duke}$ ou $\hat{p}_{EUA}$) a estimativa agrupada está mais próxima? Por quê?

Data	Duke	EUA
Muito	69	454
Nada	36	226
Total	105	680
$\hat{p}$	0.657	0.668

$$\hat{p} = \frac{\#\ de\ sucessos_1 + \#\ de\ sucessos_2}{n_1 + n_2}$$

$$= \frac{69 + 454}{105 + 680} = \frac{523}{785}$$

Prática

Calcular a estimativa proporção combinada de estudantes da Duke e americanos que se incomodariam muito com o derretimento da calota de gelo do polo norte. De qual proporção amostrada ($\hat{p}_{Duke}$ ou $\hat{p}_{EUA}$) a estimativa agrupada está mais próxima? Por quê?

Data	Duke	EUA
Muito	69	454
Nada	36	226
Total	105	680
$\hat{p}$	0.657	0.668

$$\begin{aligned} \hat{p} &= \frac{\#\ de\ sucessos_1 + \#\ de\ sucessos_2}{n_1 + n_2} \\ &= \frac{69 + 454}{105 + 680} = \frac{523}{785} = 0.666 \end{aligned}$$

Prática

Esses dados sugerem que a proporção de todos os alunos da Duke que seriam muito incomodados pelo derretimento da calota de gelo do norte difere da proporção de todos os americanos que o fazem? Calcule a estatística de teste, o valor p e interprete sua conclusão no contexto dos dados.

Dados	Duke	US
Muito	69	454
Nada	36	226
Total	105	680
$\hat{p}$	0.657	0.668

Prática

Esses dados sugerem que a proporção de todos os alunos da Duke que seriam muito incomodados pelo derretimento da calota de gelo do norte difere da proporção de todos os americanos que o fazem? Calcule a estatística de teste, o valor p e interprete sua conclusão no contexto dos dados.

Dados	Duke	US
Muito	69	454
Nada	36	226
Total	105	680
$\hat{p}$	0.657	0.668

Prática

$$Z = \frac{(\hat{p}_{Duke} - \hat{p}_{EUA})}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_{Duke}} + \frac{\hat{p}(1-\hat{p})}{n_{EUA}}}}$$

Prática

$$
\begin{aligned}
Z &= \frac{(\hat{p}_{Duke} - \hat{p}_{EUA})}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_{Duke}} + \frac{\hat{p}(1-\hat{p})}{n_{EUA}}}} \\
&= \frac{(0.657 - 0.668)}{\sqrt{\frac{0.666 \times 0.334}{105} + \frac{0.666 \times 0.334}{680}}} =
\end{aligned}
$$

Prática

$$
\begin{aligned}
Z &= \frac{(\hat{p}_{Duke} - \hat{p}_{EUA})}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_{Duke}} + \frac{\hat{p}(1-\hat{p})}{n_{EUA}}}} \\
&= \frac{(0.657 - 0.668)}{\sqrt{\frac{0.666 \times 0.334}{105} + \frac{0.666 \times 0.334}{680}}} = \frac{-0.011}{0.0495}
\end{aligned}
$$

Prática

$$
\begin{aligned}
Z &= \frac{(\hat{p}_{Duke} - \hat{p}_{EUA})}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_{Duke}} + \frac{\hat{p}(1-\hat{p})}{n_{EUA}}}} \\
&= \frac{(0.657 - 0.668)}{\sqrt{\frac{0.666\times 0.334}{105} + \frac{0.666\times 0.334}{680}}} = \frac{-0.011}{0.0495} = -0.22
\end{aligned}
$$

Prática

$$
\begin{aligned}
Z &= \frac{(\hat{p}_{Duke} - \hat{p}_{EUA})}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_{Duke}} + \frac{\hat{p}(1-\hat{p})}{n_{EUA}}}} \\
&= \frac{(0.657 - 0.668)}{\sqrt{\frac{0.666 \times 0.334}{105} + \frac{0.666 \times 0.334}{680}}} = \frac{-0.011}{0.0495} = -0.22 \\
p - valor &= 2 \times P(Z < -0.22)
\end{aligned}
$$

Prática

$$
\begin{aligned}
Z &= \frac{(\hat{p}_{Duke} - \hat{p}_{EUA})}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_{Duke}} + \frac{\hat{p}(1-\hat{p})}{n_{EUA}}}} \\
&= \frac{(0.657 - 0.668)}{\sqrt{\frac{0.666\times 0.334}{105} + \frac{0.666\times 0.334}{680}}} = \frac{-0.011}{0.0495} = -0.22 \\
p - valor &= 2 \times P(Z < -0.22) = 2 \times 0.41 = 0.82
\end{aligned}
$$

Recapitulação - comparando duas proporções

- Parâmetro de população: $(p_1 - p_2)$, estimativa pontual: $(\hat{p}_1 - \hat{p}_2)$

Recapitulação - comparando duas proporções

- Parâmetro de população: $(p_1 - p_2)$, estimativa pontual: $(\hat{p}_1 - \hat{p}_2)$
- Condições:

Recapitulação - comparando duas proporções

- Parâmetro de população: $(p_1 - p_2)$, estimativa pontual: $(\hat{p}_1 - \hat{p}_2)$
- Condições:
  - independência dentro dos grupos
    - amostra aleatória e condição de 10% atendidas para ambos os grupos
  - independência entre grupos
  - pelo menos 10 sucessos e falhas em cada grupo
    - se não → aleatorizar (Seção 6.4)

Recapitulação - comparando duas proporções

- Parâmetro de população: $(p_1 - p_2)$, estimativa pontual: $(\hat{p}_1 - \hat{p}_2)$
- Condições:
  - independência dentro dos grupos
    - amostra aleatória e condição de 10% atendidas para ambos os grupos
  - independência entre grupos
  - pelo menos 10 sucessos e falhas em cada grupo
    - se não $\rightarrow$ aleatorizar (Seção 6.4)
- $SE_{(\hat{p}_1 - \hat{p}_2)} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$
  - para IC: usar $\hat{p}_1$ e $\hat{p}_2$
  - para TH:
    - quando $H_0 : p_1 = p_2$: usar $\hat{p}_{agrupado} = \frac{\#\ suc_1 + \#suc_2}{n_1 + n_2}$
    - quando $H_0 : p_1 - p_2 =$ *(algum valor diferente de 0)*: usar $\hat{p}_1$ e $\hat{p}_2$
      - isso é muito raro

Referência - cálculos de erro padrão

	uma amostra	duas amostras
média	$SE = \frac{s}{\sqrt{n}}$	$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
proporção	$SE = \sqrt{\frac{p(1-p)}{n}}$	$SE = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

Referência - cálculos de erro padrão

	uma amostra	duas amostras
média	$SE = \frac{s}{\sqrt{n}}$	$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
proporção	$SE = \sqrt{\frac{p(1-p)}{n}}$	$SE = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

- Ao trabalhar com médias, é muito raro que σ seja conhecido, então geralmente usamos s.

Referência - cálculos de erro padrão

	uma amostra	duas amostras
média	$SE = \frac{s}{\sqrt{n}}$	$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
proporção	$SE = \sqrt{\frac{p(1-p)}{n}}$	$SE = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

- Ao trabalhar com médias, é muito raro que σ seja conhecido, então geralmente usamos s.
- Ao trabalhar com proporções,
  - se fazemos um teste de hipótese, p vem da hipótese nula
  - se construimos um intervalo de confiança, use $\hat{p}$

6.3. Teste qui-quadrado de *Goodness of fit*

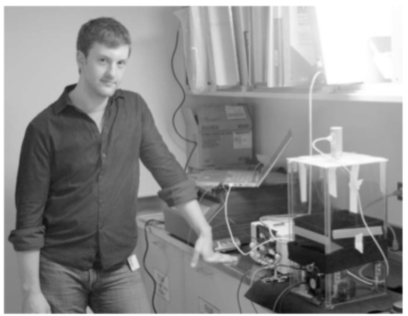
Dados de Weldon

- Walter Frank Raphael Weldon (1860 - 1906), foi um biólogo inglês evolucionário e um dos fundadores da biometria. Além disso, ele foi o editor fundador da revista *Biometrika*, com Francis Galton e Karl Pearson.
- Em 1894, ele jogou 12 dados 26.306 vezes e registrou a quatidade de vezes que sairam os número 5 e 6 (que ele considerou como sucesso).
    - Observou-se que os números 5 ou 6 ocorreram com mais frequência do que o esperado, e Pearson levantou a hipótese de que isso ocorreu provavelmente devido à construção dos dados. Os dados mais baratos têm pips (aqueles pontinhos pretos) ocos e, como os lados opostos aumentam para 7, a face com 6 pips é mais leve do que a face oposta, que tem apenas 1 pip.

Dados de Labby

- Em 2009, Zacariah Labby (U de Chicago), repetiu o experimento de Weldon usando uma máquina de contagem de lançadores de dados caseiros.

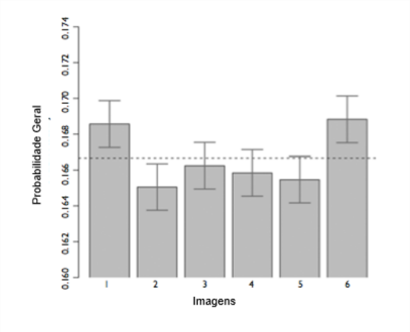
  http:// www.youtube.com/ watch?v= 95EErdouO2w



- O processo de criação de imagens demorava cerca de 20 segundos por rolagem.
  - Cada dia havia ∼ 150 imagens para processar manualmente.
  - Nesse ritmo, a experiência de Weldon foi repetida em pouco mais de seis dias inteiros.
  - Leitura recomendada: *http:// galton.uchicago.edu/ about/ docs/ labby09dice.pdf*

Dados de Labby (cont.)

- Labby não observou o mesmo fenômeno que Weldon (maior frequência dos números de 5 e 6).
- A automação permitiu que Labby coletasse mais dados do que Weldon em 1894, e, em vez de registrar "sucessos"e "falhas", Labby registrava o número individual de pips em cada dado.



Contagens esperadas

Labby rolou 12 dados 26.306 vezes. Se cada lado tiver a mesma probabilidade de aparecer, quantos 1s, 2s, $\cdots$, 6s ele esperaria observar?

(a) $\frac{1}{6}$

(b) $\frac{12}{6}$

(c) $\frac{26{,}306}{6}$

(d) $\frac{12 \times 26{,}306}{6}$

Contagens esperadas

Labby rolou 12 dados 26.306 vezes. Se cada lado tiver a mesma probabilidade de aparecer, quantos 1s, 2s, $\cdots$, 6s ele esperaria observar?

(a) $\frac{1}{6}$

(b) $\frac{12}{6}$

(c) $\frac{26{,}306}{6}$

(d) $\frac{12 \times 26{,}306}{6} = 52{,}612$

Sumarizando os resultados de Labby

A tabela abaixo mostra as contagens observadas e esperadas da experiência de Labby.

Resultado	Observado	Esperado
1	53,222	52,612
2	52,118	52,612
3	52,465	52,612
4	52,338	52,612
5	52,244	52,612
6	53,285	52,612
Total	315,672	315,672

Sumarizando os resultados de Labby

A tabela abaixo mostra as contagens observadas e esperadas da experiência de Labby.

Resultado	Observado	Esperado
1	53,222	52,612
2	52,118	52,612
3	52,465	52,612
4	52,338	52,612
5	52,244	52,612
6	53,285	52,612
Total	315,672	315,672

Por que, para todos os resultados, as contagens esperadas são as mesmas mas as contagens observadas são diferentes? À primeira vista, parece haver uma inconsistência entre as contagens observadas e esperadas?

Definindo as hipóteses

Esses dados fornecem evidências convincentes de uma inconsistência entre as contagens observadas e esperadas?

Definindo as hipóteses

Esses dados fornecem evidências convincentes de uma inconsistência entre as contagens observadas e esperadas?

H_0: Não há inconsistência entre as contagens observadas e as esperadas. *As contagens observadas seguem a mesma distribuição que as contagens esperadas.*

Definindo as hipóteses

Esses dados fornecem evidências convincentes de uma inconsistência entre as contagens observadas e esperadas?

H_0: Não há inconsistência entre as contagens observadas e as esperadas. *As contagens observadas seguem a mesma distribuição que as contagens esperadas.*

H_A: Existe uma inconsistência entre as contagens observadas e as esperadas. *As contagens observadas não seguem a mesma distribuição das contagens esperadas.* Existe um viés para que um lado apareça mais vezes no lançamento de um dado.

Avaliando as hipóteses

- Para avaliar essas hipóteses, quantificamos quão diferentes são as contagens observadas das contagens esperadas.

Avaliando as hipóteses

- Para avaliar essas hipóteses, quantificamos quão diferentes são as contagens observadas das contagens esperadas.
- Grandes desvios do que seria esperado com base na variação amostral por si só fornecem fortes evidências para a hipótese alternativa.

Avaliando as hipóteses

- Para avaliar essas hipóteses, quantificamos quão diferentes são as contagens observadas das contagens esperadas.
- Grandes desvios do que seria esperado com base na variação amostral por si só fornecem fortes evidências para a hipótese alternativa.
- Isso é chamado de teste de *goodness of fit*, ou *qualidade de ajuste*, pois estamos avaliando o quão bem os dados observados se ajustam à distribuição esperada.

Anatomia de uma estatística de teste

- A forma geral de uma estatística de teste é

$$\frac{\text{estimativa pontual} - \text{valor nulo}}{\text{erro padrão da estimativa pontual}}$$

Anatomia de uma estatística de teste

- A forma geral de uma estatística de teste é

$$\frac{\text{estimativa pontual} - \text{valor nulo}}{\text{erro padrão da estimativa pontual}}$$

- Esta construção é baseada em
  1. identificar a diferença entre uma estimativa pontual e um valor esperado, supondo que a hipótese nula é verdadeira, e
  2. padronizar essa diferença usando o erro padrão da estimativa pontual.

Anatomia de uma estatística de teste

- A forma geral de uma estatística de teste é

$$\frac{\text{estimativa pontual} - \text{valor nulo}}{\text{erro padrão da estimativa pontual}}$$

- Esta construção é baseada em
  1. identificar a diferença entre uma estimativa pontual e um valor esperado, supondo que a hipótese nula é verdadeira, e
  2. padronizar essa diferença usando o erro padrão da estimativa pontual.

  Essas duas ideias ajudarão na construção de uma estatística de teste apropriada para dados de contagem.

Estatística qui-quadrado

Ao lidar com contagens e investigar quão longe as contagens observadas estão das contagens esperadas, usamos uma nova estatística de teste chamada *estatística de qui-quadrado (χ^2)*.

Estatística qui-quadrado

Ao lidar com contagens e investigar quão longe as contagens observadas estão das contagens esperadas, usamos uma nova estatística de teste chamada *estatística de qui-quadrado (χ^2)*.

Estatística de χ^2

$$\chi^2 = \sum_{i=1}^{k} \frac{(O - E)^2}{E} \qquad \text{onde } k = \text{número total de células}$$

Calculando a estatística do qui-quadrado

Resultado	Observado	Esperado	$\frac{(O-E)^2}{E}$
1	53,222	52,612	$\frac{(53{,}222-52{,}612)^2}{52{,}612} = 7.07$

Calculando a estatística do qui-quadrado

Resultado	Observado	Esperado	$\frac{(O-E)^2}{E}$
1	53,222	52,612	$\frac{(53{,}222-52{,}612)^2}{52{,}612} = 7.07$
2	52,118	52,612	$\frac{(52{,}118-52{,}612)^2}{52{,}612} = 4.64$

Calculando a estatística do qui-quadrado

Resultado	Observado	Esperado	$\frac{(O-E)^2}{E}$
1	53,222	52,612	$\frac{(53{,}222-52{,}612)^2}{52{,}612} = 7.07$
2	52,118	52,612	$\frac{(52{,}118-52{,}612)^2}{52{,}612} = 4.64$
3	52,465	52,612	$\frac{(52{,}465-52{,}612)^2}{52{,}612} = 0.41$

Calculando a estatística do qui-quadrado

Resultado	Observado	Esperado	$\frac{(O-E)^2}{E}$
1	53,222	52,612	$\frac{(53{,}222-52{,}612)^2}{52{,}612} = 7.07$
2	52,118	52,612	$\frac{(52{,}118-52{,}612)^2}{52{,}612} = 4.64$
3	52,465	52,612	$\frac{(52{,}465-52{,}612)^2}{52{,}612} = 0.41$
4	52,338	52,612	$\frac{(52{,}338-52{,}612)^2}{52{,}612} = 1.43$

Calculando a estatística do qui-quadrado

Resultado	Observado	Esperado	$\frac{(O-E)^2}{E}$
1	53,222	52,612	$\frac{(53{,}222-52{,}612)^2}{52{,}612} = 7.07$
2	52,118	52,612	$\frac{(52{,}118-52{,}612)^2}{52{,}612} = 4.64$
3	52,465	52,612	$\frac{(52{,}465-52{,}612)^2}{52{,}612} = 0.41$
4	52,338	52,612	$\frac{(52{,}338-52{,}612)^2}{52{,}612} = 1.43$
5	52,244	52,612	$\frac{(52{,}244-52{,}612)^2}{52{,}612} = 2.57$

Calculando a estatística do qui-quadrado

Resultado	Observado	Esperado	$\frac{(O-E)^2}{E}$
1	53,222	52,612	$\frac{(53{,}222-52{,}612)^2}{52{,}612} = 7.07$
2	52,118	52,612	$\frac{(52{,}118-52{,}612)^2}{52{,}612} = 4.64$
3	52,465	52,612	$\frac{(52{,}465-52{,}612)^2}{52{,}612} = 0.41$
4	52,338	52,612	$\frac{(52{,}338-52{,}612)^2}{52{,}612} = 1.43$
5	52,244	52,612	$\frac{(52{,}244-52{,}612)^2}{52{,}612} = 2.57$
6	53,285	52,612	$\frac{(53{,}285-52{,}612)^2}{52{,}612} = 8.61$

Calculando a estatística do qui-quadrado

Resultado	Observado	Esperado	$\frac{(O-E)^2}{E}$
1	53,222	52,612	$\frac{(53{,}222-52{,}612)^2}{52{,}612} = 7.07$
2	52,118	52,612	$\frac{(52{,}118-52{,}612)^2}{52{,}612} = 4.64$
3	52,465	52,612	$\frac{(52{,}465-52{,}612)^2}{52{,}612} = 0.41$
4	52,338	52,612	$\frac{(52{,}338-52{,}612)^2}{52{,}612} = 1.43$
5	52,244	52,612	$\frac{(52{,}244-52{,}612)^2}{52{,}612} = 2.57$
6	53,285	52,612	$\frac{(53{,}285-52{,}612)^2}{52{,}612} = 8.61$
Total	315,672	315,672	24.73

Por que qui-quadrado?

Tomar o quadrado da diferença entre o resultado observado e o esperado permite duas coisas duas coisas:

Por que qui-quadrado?

Tomar o quadrado da diferença entre o resultado observado e o esperado permite duas coisas duas coisas:

- Qualquer diferença padronizada que seja quadrada será agora positiva.

Por que qui-quadrado?

Tomar o quadrado da diferença entre o resultado observado e o esperado permite duas coisas duas coisas:

- Qualquer diferença padronizada que seja quadrada será agora positiva.
- Diferenças que já pareciam incomuns se tornarão muito maiores depois de se tomar o quadrado.

Por que qui-quadrado?

Tomar o quadrado da diferença entre o resultado observado e o esperado permite duas coisas duas coisas:

- Qualquer diferença padronizada que seja quadrada será agora positiva.
- Diferenças que já pareciam incomuns se tornarão muito maiores depois de se tomar o quadrado.

Quando já vimos isso antes?

A distribuição de qui-quadrado

- Para determinar se a estatística χ^2 que calculamos é considerada excepcionalmente alta ou não, precisamos primeiro descrever sua distribuição.

A distribuição de qui-quadrado

- Para determinar se a estatística χ^2 que calculamos é considerada excepcionalmente alta ou não, precisamos primeiro descrever sua distribuição.
- A distribuição qui-quadrado tem apenas um parâmetro chamado *graus de liberdade (df)*, que influencia a forma, o centro e a largura da distribuição.

A distribuição de qui-quadrado

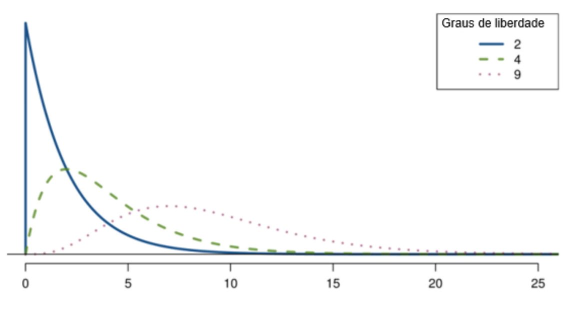
- Para determinar se a estatística χ^2 que calculamos é considerada excepcionalmente alta ou não, precisamos primeiro descrever sua distribuição.
- A distribuição qui-quadrado tem apenas um parâmetro chamado *graus de liberdade (df)*, que influencia a forma, o centro e a largura da distribuição.

Lembre-se que até agora, vimos três outras distribuições contínuas:

- Distribuição normal: unimodal e simétrica com dois parâmetros (média e desvio padrão).
- Distribuição T: unimodal e simétrica com um parâmetro (graus de liberdade).
- Distribuição F: unimodal e inclinada à direita com dois parâmetros (graus de liberdade ou numerador, que é a variância entre grupos, e denominador, que é a varição dentro do grupo).

Prática

Qual das seguintes afirmações é falsa?



À medida que df aumenta,

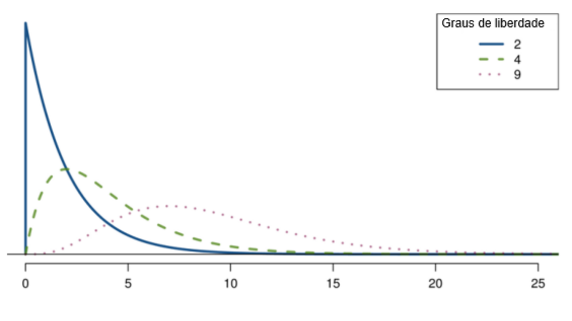
(a) o centro da distribuição χ^2 aumenta também.

(b) a variabilidade da distribuição χ^2 aumenta também.

(c) a forma da distribuição χ^2 torna-se mais assimétrica (menos parecida com uma normal).

Prática

Qual das seguintes afirmações é falsa?



À medida que df aumenta,

(a) o centro da distribuição χ^2 aumenta também.

(b) a variabilidade da distribuição χ^2 aumenta também.

(c) *a forma da distribuição χ^2 torna-se mais assimétrica (menos parecida com uma normal).*

Encontrando a área sob a curva da distribuição qui-quadrado

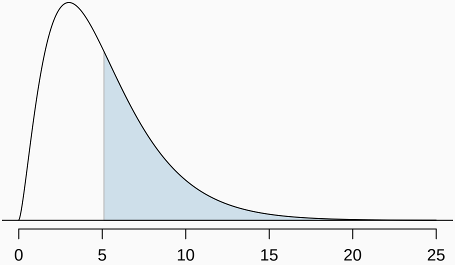
- P-valor = área da cauda sob a distribuição qui-quadrado.

Encontrando a área sob a curva da distribuição qui-quadrado

- P-valor = área da cauda sob a distribuição qui-quadrado.
- Para isso, podemos usar uma linguagem de programação, ou uma *tabela de probabilidades qui-quadrado*.

Encontrando a área sob a curva da distribuição qui-quadrado

- P-valor = área da cauda sob a distribuição qui-quadrado.
- Para isso, podemos usar uma linguagem de programação, ou uma *tabela de probabilidades qui-quadrado*.
- Esta tabela funciona como a tabela t, mas fornece apenas os valores de cauda superiores.

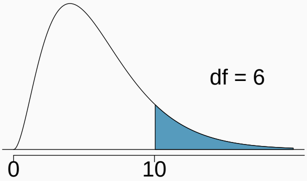


Encontrando a área sob a curva da distribuição qui-quadrado

Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001
df	1	1.07	1.64	2.71	3.84	5.41	6.63	7.88	10.83
	2	2.41	3.22	4.61	5.99	7.82	9.21	10.60	13.82
	3	3.66	4.64	6.25	7.81	9.84	11.34	12.84	16.27
	4	4.88	5.99	7.78	9.49	11.67	13.28	14.86	18.47
	5	6.06	7.29	9.24	11.07	13.39	15.09	16.75	20.52
	6	7.23	8.56	10.64	12.59	15.03	16.81	18.55	22.46
	7	8.38	9.80	12.02	14.07	16.62	18.48	20.28	24.32
	...								

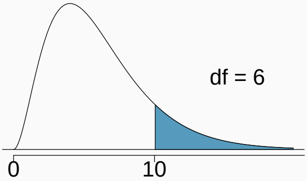
Encontrando a área sob a curva da distribuição qui-quadrado

Estimar a área sombreada sob a curva da distribuição qui-quadrado com $df = 6$.



Encontrando a área sob a curva da distribuição qui-quadrado

Estimar a área sombreada sob a curva da distribuição qui-quadrado com $df = 6$.



$P(\chi^2_{df=6} > 10)$
está entre 0.1 e 0.2

Encontrando a área sob a curva da distribuição qui-quadrado

Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001
df	1	1.07	1.64	2.71	3.84	5.41	6.63	7.88	10.83
	2	2.41	3.22	4.61	5.99	7.82	9.21	10.60	13.82
	3	3.66	4.64	6.25	7.81	9.84	11.34	12.84	16.27
	4	4.88	5.99	7.78	9.49	11.67	13.28	14.86	18.47
	5	6.06	7.29	9.24	11.07	13.39	15.09	16.75	20.52
	6	7.23	8.56	10.64	12.59	15.03	16.81	18.55	22.46
	7	8.38	9.80	12.02	14.07	16.62	18.48	20.28	24.32

Encontrando a área sob a curva da distribuição qui-quadrado

Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001
df	1	1.07	1.64	2.71	3.84	5.41	6.63	7.88	10.83
	2	2.41	3.22	4.61	5.99	7.82	9.21	10.60	13.82
	3	3.66	4.64	6.25	7.81	9.84	11.34	12.84	16.27
	4	4.88	5.99	7.78	9.49	11.67	13.28	14.86	18.47
	5	6.06	7.29	9.24	11.07	13.39	15.09	16.75	20.52
	6	7.23	8.56	10.64	12.59	15.03	16.81	18.55	22.46
	7	8.38	9.80	12.02	14.07	16.62	18.48	20.28	24.32

Encontrando a área sob a curva da distribuição qui-quadrado

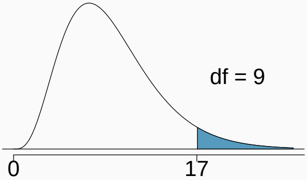
Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001
df	1	1.07	1.64	2.71	3.84	5.41	6.63	7.88	10.83
	2	2.41	3.22	4.61	5.99	7.82	9.21	10.60	13.82
	3	3.66	4.64	6.25	7.81	9.84	11.34	12.84	16.27
	4	4.88	5.99	7.78	9.49	11.67	13.28	14.86	18.47
	5	6.06	7.29	9.24	11.07	13.39	15.09	16.75	20.52
	6	7.23	*8.56*	*10.64*	12.59	15.03	16.81	18.55	22.46
	7	8.38	9.80	12.02	14.07	16.62	18.48	20.28	24.32

Encontrando a área sob a curva da distribuição qui-quadrado

Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001
df	1	1.07	1.64	2.71	3.84	5.41	6.63	7.88	10.83
	2	2.41	3.22	4.61	5.99	7.82	9.21	10.60	13.82
	3	3.66	4.64	6.25	7.81	9.84	11.34	12.84	16.27
	4	4.88	5.99	7.78	9.49	11.67	13.28	14.86	18.47
	5	6.06	7.29	9.24	11.07	13.39	15.09	16.75	20.52
	6	7.23	*8.56*	*10.64*	12.59	15.03	16.81	18.55	22.46
	7	8.38	9.80	12.02	14.07	16.62	18.48	20.28	24.32

Encontrando a área sob a curva da distribuição qui-quadrado

Estime a área sombreada (acima de 17) sob a curva χ^2 com $df = 9$.



(a) 0.05

(b) 0.02

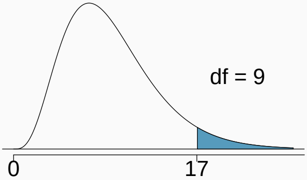
(c) entre 0.02 e 0.05

(d) entre 0.05 e 0.1

(e) entre 0.01 e 0.02

Encontrando a área sob a curva da distribuição qui-quadrado

Estime a área sombreada (acima de 17) sob a curva χ^2 com $df = 9$.



(a) 0.05

(b) 0.02

(c) *entre 0.02 e 0.05*

(d) entre 0.05 e 0.1

(e) entre 0.01 e 0.02

Encontrando a área sob a curva da distribuição qui-quadrado

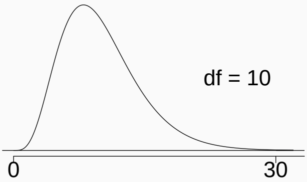
Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001
df	7	8.38	9.80	12.02	14.07	16.62	18.48	20.28	24.32
	8	9.52	11.03	13.36	15.51	18.17	20.09	21.95	26.12
	9	10.66	12.24	14.68	16.92	19.68	21.67	23.59	27.88
	10	11.78	13.44	15.99	18.31	21.16	23.21	25.19	29.59
	11	12.90	14.63	17.28	19.68	22.62	24.72	26.76	31.26

Encontrando a área sob a curva da distribuição qui-quadrado

Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001
df	7	8.38	9.80	12.02	14.07	16.62	18.48	20.28	24.32
	8	9.52	11.03	13.36	15.51	18.17	20.09	21.95	26.12
	9	10.66	12.24	14.68	*16.92*	*19.68*	21.67	23.59	27.88
	10	11.78	13.44	15.99	18.31	21.16	23.21	25.19	29.59
	11	12.90	14.63	17.28	19.68	22.62	24.72	26.76	31.26

Encontrando a área sob a curva da distribuição qui-quadrado

Estime a área sombreada (acima de 30) sob a curva χ^2 com $df = 10$.



(a) maior que 0.3

(b) entre 0.005 e 0.001

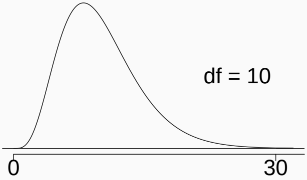
(c) menor que 0.001

(d) maior que 0.001

(e) não posso dizer usando esta tabela

Encontrando a área sob a curva da distribuição qui-quadrado

Estime a área sombreada (acima de 30) sob a curva χ^2 com $df = 10$.



(a) maior que 0.3

(b) entre 0.005 e 0.001

(c) *menor que 0.001*

(d) maior que 0.001

(e) não posso dizer usando esta tabela

Encontrando a área sob a curva da distribuição qui-quadrado

Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001
df	7	8.38	9.80	12.02	14.07	16.62	18.48	20.28	24.32
	8	9.52	11.03	13.36	15.51	18.17	20.09	21.95	26.12
	9	10.66	12.24	14.68	16.92	19.68	21.67	23.59	27.88
	10	11.78	13.44	15.99	18.31	21.16	23.21	25.19	29.59
	11	12.90	14.63	17.28	19.68	22.62	24.72	26.76	31.26

Encontrando a área sob a curva da distribuição qui-quadrado

Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	*0.001*	→
df	7	8.38	9.80	12.02	14.07	16.62	18.48	20.28	24.32	
	8	9.52	11.03	13.36	15.51	18.17	20.09	21.95	26.12	
	9	10.66	12.24	14.68	16.92	19.68	21.67	23.59	27.88	
	10	11.78	13.44	15.99	18.31	21.16	23.21	25.19	*29.59*	→
	11	12.90	14.63	17.28	19.68	22.62	24.72	26.76	31.26	

Encontrando a área sob a curva da distribuição qui-quadrado usando computação

- Embora as tabelas de probabilidades sejam muito úteis para entender como as distribuições de probabilidade funcionam e fornecer uma referência rápida quando os recursos computacionais não estão disponíveis, elas são um pouco arcaicas.

Encontrando a área sob a curva da distribuição qui-quadrado usando computação

- Embora as tabelas de probabilidades sejam muito úteis para entender como as distribuições de probabilidade funcionam e fornecer uma referência rápida quando os recursos computacionais não estão disponíveis, elas são um pouco arcaicas.
- Usando R:

```
pchisq(q = 30, df = 10, lower.tail = FALSE)
# 0.0008566412
```

Encontrando a área sob a curva da distribuição qui-quadrado usando computação

- Embora as tabelas de probabilidades sejam muito úteis para entender como as distribuições de probabilidade funcionam e fornecer uma referência rápida quando os recursos computacionais não estão disponíveis, elas são um pouco arcaicas.
- Usando R:

```
pchisq(q = 30, df = 10, lower.tail = FALSE)
# 0.0008566412
```

- Usando um applet da web:
  http://bitly.com/dist_calc

De volta aos dados de Labby

- A questão de pesquisa era: Esses dados fornecem evidências convincentes de uma inconsistência entre as contagens observadas e as esperadas?

De volta aos dados de Labby

- A questão de pesquisa era: Esses dados fornecem evidências convincentes de uma inconsistência entre as contagens observadas e as esperadas?
- As hipóteses eram:

  H_0: Não há inconsistência entre as contagens observadas e as esperadas. As contagens observadas seguem a mesma distribuição que as contagens esperadas.

  H_A: Existe uma inconsistência entre as contagens observadas e as esperadas. As contagens observadas *não* seguem a mesma distribuição das contagens esperadas. Existe um viés a respeito do lado que aparecerá no lançamento de um dado.

De volta aos dados de Labby

- A questão de pesquisa era: Esses dados fornecem evidências convincentes de uma inconsistência entre as contagens observadas e as esperadas?
- As hipóteses eram:

  H_0: Não há inconsistência entre as contagens observadas e as esperadas. As contagens observadas seguem a mesma distribuição que as contagens esperadas.

  H_A: Existe uma inconsistência entre as contagens observadas e as esperadas. As contagens observadas *não* seguem a mesma distribuição das contagens esperadas. Existe um viés a respeito do lado que aparecerá no lançamento de um dado.
- Nós calculamos a estatística de teste de $\chi^2 = 24.67$.

De volta aos dados de Labby

- A questão de pesquisa era: Esses dados fornecem evidências convincentes de uma inconsistência entre as contagens observadas e as esperadas?
- As hipóteses eram:

  H_0: Não há inconsistência entre as contagens observadas e as esperadas. As contagens observadas seguem a mesma distribuição que as contagens esperadas.

  H_A: Existe uma inconsistência entre as contagens observadas e as esperadas. As contagens observadas *não* seguem a mesma distribuição das contagens esperadas. Existe um viés a respeito do lado que aparecerá no lançamento de um dado.
- Nós calculamos a estatística de teste de $\chi^2 = 24.67$.
- Tudo o que precisamos é do df e podemos calcular a área final (o p-valor) e tomar uma decisão sobre as hipóteses.

Graus de liberdade de um teste de ajuste

- Ao realizar um teste de qualidade do ajuste para avaliar quão bem os dados observados seguem uma distribuição esperada, os graus de liberdade são calculados como o número de células (k) menos 1.

$$df = k - 1$$

Graus de liberdade de um teste de ajuste

- Ao realizar um teste de qualidade do ajuste para avaliar quão bem os dados observados seguem uma distribuição esperada, os graus de liberdade são calculados como o número de células (k) menos 1.

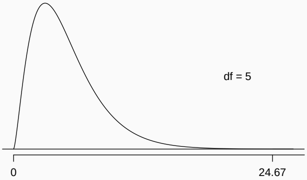
$$df = k - 1$$

- Para resultados de dados, $k = 6$, portanto

$$df = 6 - 1 = 5$$

Encontrar o p-valor de um teste do qui-quadrado

O *p-valor* de um teste qui-quadrado é definido como a *área de cauda acima da estatística de teste calculada*.



p-valor = $P(\chi^2_{df=5} > 24.67)$
é menos do que 0.001

Encontrar o p-valor de um teste do qui-quadrado

Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001	$\rightarrow$
df	1	1.07	1.64	2.71	3.84	5.41	6.63	7.88	10.83	
	2	2.41	3.22	4.61	5.99	7.82	9.21	10.60	13.82	
	3	3.66	4.64	6.25	7.81	9.84	11.34	12.84	16.27	
	4	4.88	5.99	7.78	9.49	11.67	13.28	14.86	18.47	
	5	6.06	7.29	9.24	11.07	13.39	15.09	16.75	20.52	$\rightarrow$

Conclusão do teste de hipóteses

Calculamos um p-valor menor que 0.001. Ao nível de significância de 5%, qual é a conclusão do teste de hipótese?

(a) Rejeitar H_0, os dados fornecem evidências convincentes de que os dados são honestos.

(b) Rejeitar H_0, os dados fornecem evidências convincentes de que os dados são tendenciosos.

(c) Não rejeitar H_0, os dados fornecem evidências convincentes de que os dados são honestos.

(d) Não rejeitar H_0, os dados fornecem evidências convincentes de que os dados são tendenciosos.

Conclusão do teste de hipóteses

Calculamos um p-valor menor que 0.001. Ao nível de significância de 5%, qual é a conclusão do teste de hipótese?

(a) Rejeitar H_0, os dados fornecem evidências convincentes de que os dados são honestos.

(b) *Rejeitar H_0, os dados fornecem evidências convincentes de que os dados são tendenciosos.*

(c) Não rejeitar H_0, os dados fornecem evidências convincentes de que os dados são honestos.

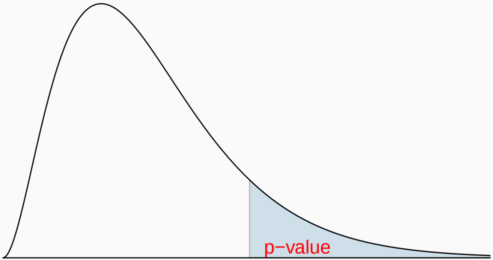
(d) Não rejeitar H_0, os dados fornecem evidências convincentes de que os dados são tendenciosos.

Acontece que...

- O eixo 1-6 é consistentemente mais curto que os outros dois (2-5 e 3-4), suportando assim a hipótese de que as faces com 1 e 6 são maiores que as outras faces.
- A alegação de Pearson de que 5s e 6s aparecem com mais frequência devido aos pips esculpidos não é suportada por esses dados.
- Os dados usados nos cassinos têm faces niveladas, onde os pontos são preenchidos com um plástico da mesma densidade do material circundante e são precisamente balanceados.

Recapitulando: p-valor do teste qui-quadrado

- O p-valor do teste qui-quadrado é definido como a área de cauda *acima* da estatística de teste calculada.
- Isso ocorre porque a estatística de teste é sempre positiva e uma estatística de teste mais alta significa um desvio maior da hipótese nula.



Condições para o teste do qui-quadrado

1. *Independência:* Cada caso que contribui com uma contagem para a tabela deve ser independente de todos os outros casos na tabela.

Condições para o teste do qui-quadrado

1. *Independência:* Cada caso que contribui com uma contagem para a tabela deve ser independente de todos os outros casos na tabela.
2. *Tamanho da amostra:* Cada cenário específico (ou seja, cada célula) deve ter pelo menos 5 casos *esperados*.

Condições para o teste do qui-quadrado

1. *Independência:* Cada caso que contribui com uma contagem para a tabela deve ser independente de todos os outros casos na tabela.
2. *Tamanho da amostra:* Cada cenário específico (ou seja, cada célula) deve ter pelo menos 5 casos *esperados*.
3. *$df > 1$:* Graus de liberdade devem ser maiores que 1.

Condições para o teste do qui-quadrado

1. *Independência:* Cada caso que contribui com uma contagem para a tabela deve ser independente de todos os outros casos na tabela.
2. *Tamanho da amostra:* Cada cenário específico (ou seja, cada célula) deve ter pelo menos 5 casos *esperados*.
3. *df > 1:* Graus de liberdade devem ser maiores que 1.

Deixar de verificar as condições pode afetar as taxas de erro do teste.

Eleição de 2009 no Irã

Falou-se muito sobre fraude eleitoral nas eleições de 2009 no Irã. Vamos comparar os dados de uma pesquisa realizada antes da eleição (dados observados) com os votos informados na eleição para ver se os dois seguem a mesma distribuição.

Candidato	# observado de votos na pesquisa	% relatada de votos na eleição
(1) Ahmedinajad	338	63.29%
(2) Mousavi	136	34.10%
(3) Candidatos menores	30	2.61%
Total	504	100%

Eleição de 2009 no Irã

Falou-se muito sobre fraude eleitoral nas eleições de 2009 no Irã. Vamos comparar os dados de uma pesquisa realizada antes da eleição (dados observados) com os votos informados na eleição para ver se os dois seguem a mesma distribuição.

Candidato	# observado de votos na pesquisa	% relatada de votos na eleição
(1) Ahmedinajad	338	63.29%
(2) Mousavi	136	34.10%
(3) Candidatos menores	30	2.61%
Total	504	100%
	↓	↓
	observado	*esperado distribuição*

Hipóteses

Quais são as hipóteses para testar se as distribuições de votos informados e observados são diferentes?

Hipóteses

Quais são as hipóteses para testar se as distribuições de votos informados e observados são diferentes?

H_0: *As contagens observadas da pesquisa seguem a mesma distribuição dos votos relatados.*

H_A: *As contagens observadas da pesquisa não seguem a mesma distribuição que os votos relatados.*

Cálculo da estatística de teste

Candidato	# observado de votos na pesquisa	% relatada de votos na eleição	# esperado de votos na pesquisa
(1) Ahmedinajad	338	63.29%	$504 \times 0.6329 = 319$
(2) Mousavi	136	34.10%	$504 \times 0.3410 = 172$
(3) Candidatos menores	30	2.61%	$504 \times 0.0261 = 13$
Total	504	100%	504

Cálculo da estatística de teste

Candidato	# observado de votos na pesquisa	% relatada de votos na eleição	# esperado de votos na pesquisa
(1) Ahmedinajad	338	63.29%	$504 \times 0.6329 = 319$
(2) Mousavi	136	34.10%	$504 \times 0.3410 = 172$
(3) Candidatos menores	30	2.61%	$504 \times 0.0261 = 13$
Total	504	100%	504

$$\frac{(O_1 - E_1)^2}{E_1} = \frac{(338 - 319)^2}{319} = 1.13$$

Cálculo da estatística de teste

Candidato	# observado de votos na pesquisa	% relatada de votos na eleição	# esperado de votos na pesquisa
(1) Ahmedinajad	338	63.29%	504 × 0.6329 = 319
(2) Mousavi	136	34.10%	504 × 0.3410 = 172
(3) Candidatos menores	30	2.61%	504 × 0.0261 = 13
Total	504	100%	504

$$\frac{(O_1 - E_1)^2}{E_1} = \frac{(338 - 319)^2}{319} = 1.13$$

$$\frac{(O_2 - E_2)^2}{E_2} = \frac{(136 - 172)^2}{172} = 7.53$$

Cálculo da estatística de teste

Candidato	# observado de votos na pesquisa	% relatada de votos na eleição	# esperado de votos na pesquisa
(1) Ahmedinajad	338	63.29%	504 × 0.6329 = 319
(2) Mousavi	136	34.10%	504 × 0.3410 = 172
(3) Candidatos menores	30	2.61%	504 × 0.0261 = 13
Total	504	100%	504

$$\frac{(O_1 - E_1)^2}{E_1} = \frac{(338 - 319)^2}{319} = 1.13$$

$$\frac{(O_2 - E_2)^2}{E_2} = \frac{(136 - 172)^2}{172} = 7.53$$

$$\frac{(O_2 - E_2)^2}{E_2} = \frac{(30 - 13)^2}{13} = 22.23$$

Cálculo da estatística de teste

Candidato	# observado de votos na pesquisa	% relatada de votos na eleição	# esperado de votos na pesquisa
(1) Ahmedinajad	338	63.29%	504 × 0.6329 = 319
(2) Mousavi	136	34.10%	504 × 0.3410 = 172
(3) Candidatos menores	30	2.61%	504 × 0.0261 = 13
Total	504	100%	504

$$
\begin{aligned}
\frac{(O_1 - E_1)^2}{E_1} = \frac{(338 - 319)^2}{319} &= 1.13 \\
\frac{(O_2 - E_2)^2}{E_2} = \frac{(136 - 172)^2}{172} &= 7.53 \\
\frac{(O_2 - E_2)^2}{E_2} = \frac{(30 - 13)^2}{13} &= 22.23 \\
\chi^2_{df=3-1=2} &= 30.89
\end{aligned}
$$

Conclusão

Com base nesses cálculos, qual é a conclusão do teste de hipóteses?

(a) o p-valor é baixo, H_0 é rejeitada. As contagens observadas na pesquisa não seguem a mesma distribuição que os votos relatados.

(b) o p-valor é alto, H_0 não é rejeitada. As contagens observadas na pesquisa seguem a mesma distribuição dos votos relatados.

(c) o p-valor é baixo, H_0 é rejeitada. As contagens observadas na pesquisa seguem a mesma distribuição que os votos relatados.

(d) o p-valor é baixo, H_0 não é rejeitada. As contagens observadas na pesquisa não seguem a mesma distribuição que os votos relatados.

Conclusão

Com base nesses cálculos, qual é a conclusão do teste de hipóteses?

(a) *o p-valor é baixo, H_0 é rejeitada. As contagens observadas na pesquisa não seguem a mesma distribuição que os votos relatados.*

(b) o p-valor é alto, H_0 não é rejeitada. As contagens observadas na pesquisa seguem a mesma distribuição dos votos relatados.

(c) o p-valor é baixo, H_0 é rejeitada. As contagens observadas na pesquisa seguem a mesma distribuição que os votos relatados.

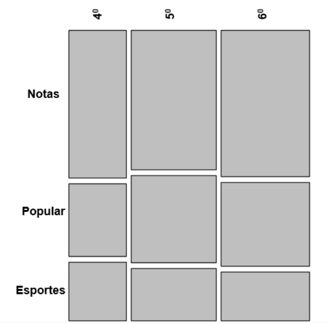
(d) o p-valor é baixo, H_0 não é rejeitada. As contagens observadas na pesquisa não seguem a mesma distribuição que os votos relatados.

6.4. Teste de independência de qui-quadrado

Crianças populares

No conjunto de dados `popular`, os alunos das 4ª a 6ª séries foram questionados se boas notas, ser atlético ou ser popular era mais importante para eles. Uma tabela de duas vias que separa os alunos por série e por escolha do fator mais importante é mostrada abaixo. Esses dados fornecem evidências para sugerir que as preferências variam de acordo com o série?

	Notas	Popular	Esportes
4^a	63	31	25
5^a	88	55	33
6^a	96	55	32



Teste de independência do qui-quadrado

- As hipóteses são:

  H_0: A série e as preferências são independentes. As preferências não variam de acordo com a série.

  H_A: A série e as preferências são dependentes. As preferências variam de acordo com a série.

Teste de independência do qui-quadrado

- As hipóteses são:

  H_0: A série e as preferências são independentes. As preferências não variam de acordo com a série.

  H_A: A série e as preferências são dependentes. As preferências variam de acordo com a série.

- A estatística de teste é calculada como

$$\chi^2_{df} = \sum_{i=1}^{k} \frac{(O - E)^2}{E} \quad \text{onde} \quad df = (R - 1) \times (C - 1),$$

  onde k é o número de células, R é o número de linhas e C é o número de colunas.

Teste de independência do qui-quadrado

- As hipóteses são:
  - H_0: A série e as preferências são independentes. As preferências não variam de acordo com a série.
  - H_A: A série e as preferências são dependentes. As preferências variam de acordo com a série.
- A estatística de teste é calculada como

$$\chi^2_{df} = \sum_{i=1}^{k} \frac{(O - E)^2}{E} \quad \text{onde} \quad df = (R - 1) \times (C - 1),$$

  onde k é o número de células, R é o número de linhas e C é o número de colunas.
- O p-valor é a área sob a curva χ^2_{df} acima da estatística de teste calculada.

Note: Calculamos df de forma diferente para tabelas de uma via e duas vias.

Contagens esperadas em tabelas de duas vias

Contagens esperadas em tabelas de duas vias

$$\text{Contagem esperada} = \frac{(\text{total da linha}) \times (\text{total da coluna})}{\text{total da tabela}}$$

Contagens esperadas em tabelas de duas vias

Contagens esperadas em tabelas de duas vias

$$\text{Contagem esperada} = \frac{(\text{total da linha}) \times (\text{total da coluna})}{\text{total da tabela}}$$

	Notas	Popular	Esportes	Total
4^a	*63*	*31*	25	119
5^a	88	55	33	176
6^a	96	55	32	183
Total	247	141	90	478

Contagens esperadas em tabelas de duas vias

Contagens esperadas em tabelas de duas vias

$$\text{Contagem esperada} = \frac{(\text{total da linha}) \times (\text{total da coluna})}{\text{total da tabela}}$$

	Notas	Popular	Esportes	Total
4^a	*63*	*31*	25	119
5^a	88	55	33	176
6^a	96	55	32	183
Total	247	141	90	478

$$E_{linha\ 1, coluna\ 1} = \frac{119 \times 247}{478} = 61$$

Contagens esperadas em tabelas de duas vias

Contagens esperadas em tabelas de duas vias

$$\text{Contagem esperada} = \frac{(\text{total da linha}) \times (\text{total da coluna})}{\text{total da tabela}}$$

	Notas	Popular	Esportes	Total
4^a	*63*	*31*	25	119
5^a	88	55	33	176
6^a	96	55	32	183
Total	247	141	90	478

$$E_{linha\ 1, coluna\ 1} = \frac{119 \times 247}{478} = 61 \qquad E_{linha\ 1, coluna\ 2} = \frac{119 \times 141}{478} = 35$$

Contagens esperadas em tabelas bidirecionais

Qual é a contagem esperada para a célula realçada?

	Notas	Popular	Esportes	Total
4^a	63	31	25	119
5^a	88	55	33	176
6^a	96	55	32	183
Total	247	141	90	478

(a) $\frac{176 \times 141}{478}$

(b) $\frac{119 \times 141}{478}$

(c) $\frac{176 \times 247}{478}$

(d) $\frac{176 \times 478}{478}$

Contagens esperadas em tabelas bidirecionais

Qual é a contagem esperada para a célula realçada?

	Notas	Popular	Esportes	Total
4^a	63	31	25	119
5^a	88	55	33	176
6^a	96	55	32	183
Total	247	141	90	478

(a) $\frac{176 \times 141}{478}$

(b) $\frac{119 \times 141}{478}$

(c) $\frac{176 \times 247}{478}$

(d) $\frac{176 \times 478}{478}$

$\rightarrow$ *52*

mais do que o # esperado de alunos do 5º ano tem preferência por ser popular

Calculando a estatística de teste em tabelas bidirecionais

As contagens esperadas são mostradas em *azul* ao lado das contagens observadas.

	Notas	Popular	Esportes	Total
4^a	63 *61*	31 *35*	25 *23*	119
5^a	88 *91*	55 *52*	33 *33*	176
6^a	96 *95*	55 *54*	32 *34*	183
Total	247	141	90	478

Calculando a estatística de teste em tabelas bidirecionais

As contagens esperadas são mostradas em *azul* ao lado das contagens observadas.

	Notas	Popular	Esportes	Total
4^a	63 *61*	31 *35*	25 *23*	119
5^a	88 *91*	55 *52*	33 *33*	176
6^a	96 *95*	55 *54*	32 *34*	183
Total	247	141	90	478

$$\chi^2 = \sum \frac{(63-61)^2}{61} + \frac{(31-35)^2}{35} + \cdots + \frac{(32-34)^2}{34} = 1.3121$$

Calculando a estatística de teste em tabelas bidirecionais

As contagens esperadas são mostradas em *azul* ao lado das contagens observadas.

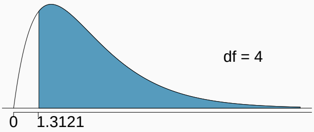
	Notas	Popular	Esportes	Total
4^a	63 *61*	31 *35*	25 *23*	119
5^a	88 *91*	55 *52*	33 *33*	176
6^a	96 *95*	55 *54*	32 *34*	183
Total	247	141	90	478

$$\begin{aligned} \chi^2 &= \sum \frac{(63-61)^2}{61} + \frac{(31-35)^2}{35} + \cdots + \frac{(32-34)^2}{34} = 1.3121 \\ df &= (R-1) \times (C-1) = (3-1) \times (3-1) = 2 \times 2 = 4 \end{aligned}$$

Calculando o p-valor

Qual dos seguintes é o p-valor correto para este teste de hipóteses?

$$\chi^2 = 1.3121 \qquad df = 4$$



(a) maior que 0.3

(b) entre 0.3 e 0.2

(c) entre 0.2 e 0.1

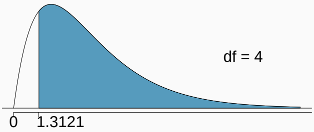
(d) entre 0.1 e 0.05

(e) menor que 0.001

Calculando o p-valor

Qual dos seguintes é o p-valor correto para este teste de hipóteses?

$$\chi^2 = 1.3121 \qquad df = 4$$



(a) *maior que 0.3*

(b) entre 0.3 e 0.2

(c) entre 0.2 e 0.1

(d) entre 0.1 e 0.05

(e) menor que 0.001

Calculando o p-valor

Cauda superior		0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.001
df	1	1.07	1.64	2.71	3.84	5.41	6.63	7.88	10.83
	2	2.41	3.22	4.61	5.99	7.82	9.21	10.60	13.82
	3	3.66	4.64	6.25	7.81	9.84	11.34	12.84	16.27
	4	4.88	5.99	7.78	9.49	11.67	13.28	14.86	18.47
	5	6.06	7.29	9.24	11.07	13.39	15.09	16.75	20.52

Conclusão

Esses dados fornecem evidências para sugerir que as preferências variam de acordo com a série?

H_0: A série e as preferências são independentes. As preferências não variam de acordo com a séries.

H_A: A série e as preferências são dependentes. As preferências variam de acordo com a série.

Conclusão

Esses dados fornecem evidências para sugerir que as preferências variam de acordo com a série?

H_0: A série e as preferências são independentes. As preferências não variam de acordo com a séries.

H_A: A série e as preferências são dependentes. As preferências variam de acordo com a série.

Como o p-valor é alto, não rejeitamos H_0. Os dados não fornecem evidências convincentes de que a série e as preferências são dependentes. Não parece que as preferências variam por série.

6.5. Inferência para uma proporção com uma amostra pequena

Preditores famosos

Antes desse cara...

Preditores famosos

Antes desse cara...

Tinha esse cara...

Paul, o polvo - psíquico?

- Paul, o polvo, previu 8 jogos da Copa do Mundo corretamente.

Paul, o polvo - psíquico?

- Paul, o polvo, previu 8 jogos da Copa do Mundo corretamente.
- Isso fornece evidências convincentes de que Paul realmente tem poderes psíquicos?

Paul, o polvo - psíquico?

- Paul, o polvo, previu 8 jogos da Copa do Mundo corretamente.
- Isso fornece evidências convincentes de que Paul realmente tem poderes psíquicos?
- Quão incomum seria se ele estivesse apenas adivinhando aleatoriamente (com 50% de chance de adivinhar corretamente)?

Paul, o polvo - psíquico?

- Paul, o polvo, previu 8 jogos da Copa do Mundo corretamente.
- Isso fornece evidências convincentes de que Paul realmente tem poderes psíquicos?
- Quão incomum seria se ele estivesse apenas adivinhando aleatoriamente (com 50% de chance de adivinhar corretamente)?
- Hipóteses:

$$H_0: \ p = 0.5$$
$$H_A: \ p > 0.5$$

Condições

1. *Independência:* Podemos supor que cada palpite é independente de outro.

Condições

1. *Independência:* Podemos supor que cada palpite é independente de outro.
2. *Tamanho da amostra:* O número de sucessos esperados é *menor que 10*.

$$8 \times 0.5 = 4$$

Condições

1. *Independência:* Podemos supor que cada palpite é independente de outro.
2. *Tamanho da amostra:* O número de sucessos esperados é *menor que 10*.

$$8 \times 0.5 = 4$$

Então, o que fazemos?

Condições

1. *Independência:* Podemos supor que cada palpite é independente de outro.
2. *Tamanho da amostra:* O número de sucessos esperados é *menor que 10*.

$$8 \times 0.5 = 4$$

Então, o que fazemos?

Como o tamanho da amostra não é grande o suficiente para usar métodos baseados em TCL, usamos um método de simulação.

Prática

Qual dos seguintes métodos é a melhor maneira de calcular o p-valor do teste de hipótese para testar se as previsões de Paul, o polvo, são excepcionalmente mais altas do que adivinhações aleatórias?

(a) Jogue uma moeda 8 vezes, registre a proporção de vezes em que todos os 8 lançamentos foram cara. Repita isso muitas vezes e calcule a proporção de simulações em que todos os 8 lançamentos foram cara.

(b) Jogue um dado 8 vezes, registre a proporção de vezes em que todos os 8 lançamentos foram 6s. Repita isso várias vezes e calcule a proporção de simulações em que todos os 8 lançamentos foram 6s.

(c) Jogue uma moeda 10.000 vezes, registre a proporção de cara. Repita isso várias vezes e calcule a proporção de simulações em que mais de 50% dos lançamentos são cara.

(d) Jogue uma moeda 10.000 vezes, calcule a proporção de cara.

Simular

Jogue uma moeda 8 vezes. Você conseguiu todas as caras?

(a) Sim
(b) Não

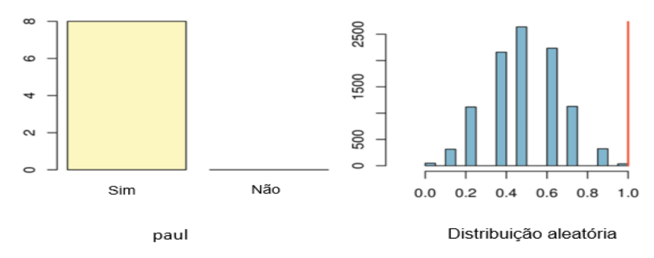
Simular

```
fonte ("http://www.openintro.org/stat/slides/inference.R")
paul = fator(c(rep("sim", 8), rep("não", 0)), nível = c("sim","não"))
inferência(paul, est = "proporção", tipo = "ht", método = "simulação",
          sucesso = "sim", nulo = 0.5, alternativo = "maior", semente = 290)
```

Simular

```
fonte ("http://www.openintro.org/stat/slides/inference.R")
paul = fator(c(rep("sim", 8), rep("não", 0)), nível = c("sim","não"))
inferência(paul, est = "proporção", tipo = "ht", método = "simulação",
          sucesso = "sim", nulo = 0.5, alternativo = "maior", semente = 290)
```

```
Proporção única -- sucesso: sim
Estatísticas resumidas: p_hat = 1 ;  n = 8
H0: p = 0.5
HA: p > 0.5
valor-p =  0.0037
```



Conclusões

Qual das seguintes opções é falsa??

(a) Se, de fato, Paul estivesse adivinhando aleatoriamente, a probabilidade de ele obter o resultado de todos os 8 jogos corretos é 0,0037.

(b) Rejeitar H_0, os dados fornecem evidências convincentes de que Paul fez mais do que adivinhar aleatoriamente.

(c) Podemos ter cometido um erro do Tipo I.

(d) A probabilidade de que Paul seja psíquico é 0,0037.

Conclusões

Qual das seguintes opções é falsa??

(a) Se, de fato, Paul estivesse adivinhando aleatoriamente, a probabilidade de ele obter o resultado de todos os 8 jogos corretos é 0,0037.

(b) Rejeitar H_0, os dados fornecem evidências convincentes de que Paul fez mais do que adivinhar aleatoriamente.

(c) Podemos ter cometido um erro do Tipo I.

(d) *A probabilidade de que Paul seja psíquico é 0,0037.*

Dorso da mão

Há um ditado que diz "conheça algo como a palma da sua mão". Descreva um experimento para testar se as pessoas conhecem *o dorso* de suas mãos.

Dorso da mão

Há um ditado que diz "conheça algo como a palma da sua mão". Descreva um experimento para testar se as pessoas conhecem *o dorso* de suas mãos.

No episódio de MythBusters, 11 de 12 pessoas adivinham as costas de suas mãos corretamente.

Hipóteses

Quais são as hipóteses para avaliar se as pessoas são capazes de reconhecer o dorso da mão a uma taxa melhor que uma adivinhação aleatória. Lembre-se, no experimento do MythBusters, havia 10 fotos para escolher, e apenas 1 estava correta.

$H_0: \ p = 0.10$ (adivinhação aleatória)

$H_A: \ p > 0.10$ (mais do que adivinhar aleatoriamente)

Condições

1. *Independência:* Podemos supor que cada pessoa que adivinhando é independente de outra.
2. *Tamanho da amostra:* O número de sucessos esperados é *menor que 10*.

$$12 \times 0.1 = 1.2$$

Então, o que fazemos?

Como o tamanho da amostra não é grande o suficiente para usar métodos baseados em TCL, usamos um método de simulação.

Esquema de simulação

Descreva como você testaria se os resultados desse experimento determinam se as pessoas são capazes de reconhecer o dorso da mão a uma taxa melhor do que uma adivinhação aleatória.

$$H_0 : p = 0.10 \qquad H_A : p > 0.10 \qquad \hat{p} = 11/12 = 0.9167$$

1. Use um dado justo de 10 lados para representar o espaço de amostragem, e chame 1 de sucesso (adivinhando corretamente), e todas os outros resultados de falhas (adivinhando incorretamente).
2. Jogue o dado 12 vezes (representando 12 pessoas no experimento), conte o número de 1s e calcule a proporção de palpites corretos em uma simulação com 12 lançamentos.
3. Repita o passo (2) muitas vezes, registrando a proporção de sucessos em uma série de 12 lançamentos do dado.
4. Crie um gráfico de pontos das proporções simuladas a partir do passo (3) e conte o número de simulações em que a proporção foi pelo menos tão alta quanto 0.9167 (a proporção observada).

Resultados simulados

- No próximo slide você pode ver os resultados de um teste de hipótese (usando apenas 100 simulações para manter as coisas simples).
- Cada ponto representa uma proporção de sucesso da simulação. Houve 25-30 simulações em que a taxa de sucesso ($\hat{p}$) foi de 10%, 40-45 simulações em que a taxa de sucesso foi ligeiramente inferior a 10%, cerca de 20 simulações em que a taxa de sucesso foi ligeiramente menor que 20% e 1 simulação em que a taxa de sucesso foi superior a 30%.
- Não há simulações em que a taxa de sucesso seja tão alta quanto a taxa de sucesso observada de 91.67%.

Resultados simulados

- Portanto, concluímos que o resultado observado é quase impossível de acontecer por acaso (p-valor = 0).
- E, portanto, esses dados sugerem que as pessoas são capazes de reconhecer as costas de suas mãos a uma taxa melhor do que adivinhar aleatoriamente.

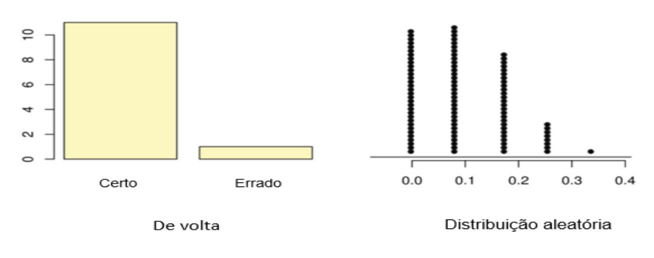
Resultados simulados

```
costas = as.factor(c(rep("correto", 11), rep("errado", 1)))
inferência(costas, est = "proporção", tipo = "ht", método = "simulação",
        sucesso = "correto", nulo = 0.1, alternativo = "maior", semente = 654, nsim = 100)
```

Resultados simulados

```
costas = as.factor(c(rep("correto", 11), rep("errado", 1)))
inferência(costas, est = "proporção", tipo = "ht", método = "simulação",
        sucesso = "correto", nulo = 0.1, alternativo = "maior", semente = 654, nsim = 100)
```

```
Proporção única - sucesso: correto
Estatísticas resumidas: p_hat = 0.9167 ;  n = 12
H0: p = 0.1
HA: p > 0.1
p-value =  0
```



6.6. Inferência para diferença entre duas proporções com uma amostra pequena

Comparando as costas da mão à palma da mão

MythBusters também pediu a essas pessoas para adivinharem as palmas das mãos. Desta vez, 7 das 12 pessoas acertaram. Os dados estão resumidos abaixo.

	Dorso	Palma	Total
Correto	11	7	18
Errado	1	5	6
Total	12	12	24

Proporção de palpites corretos

	Dorso	Palma	Total
Correto	11	7	18
Errado	1	5	6
Total	12	12	24

- Proporção de acertos no grupo do dorso: $\frac{11}{12} = 0.916$
- Proporção de correto no grupo da palme: $\frac{7}{12} = 0.583$
- Diferença: 33.3% mais acertos do grupo que avaliou o dorso das mãos.

Com base nas proporções calculadas, você acha que a chance de adivinhar corretamente o dorso da mão é diferente da palma da mão?

Hipóteses

Quais são as hipóteses para comparar se a proporção de pessoas que conseguem adivinhar corretamente o dorso de suas mãos é diferente da proporção de pessoas que conseguem adivinhar corretamente a palma de suas mãos?

H_0: $p_{dorso} = p_{palma}$

H_0: $p_{dorso} \neq p_{palma}$

Condições?

- Independência - dentro de grupos, entre grupos?
  - Dentro de cada grupo podemos supor que o palpite de um sujeito é independente do outro.
  - Entre os grupos, a independência não é satisfeita - temos as mesmas pessoas adivinhando. No entanto, vamos supor que são suposições independentes para continuar com a análise.
- Tamanho da amostra?
  - $\hat{p}_{amostra} = \frac{11+7}{12+12} = \frac{18}{24} = 0.75$.
  - Sucessos esperados no grupo do dorso: $12 \times 0.75 = 9$, falhas = 3.
  - Sucessos esperados no grupo da palma: $12 \times 0.75 = 9$, falhas = 3.
  - Como a condição Sucessos/Fracassos falha, precisamos usar a simulação para comparar as proporções.

Esquema de simulação

1. Use 24 fichas de índice, onde cada cartão representa um assunto.
2. Marque 18 das cartas como "corretas"e as 6 restantes como "erradas".
3. Embaralhe as cartas e divida em dois grupos de tamanho 12, para as costas e a palma da mão.
4. Calcule a diferença entre as proporções de "correto para o dorso"e nos decks de palma, e registre este número.
5. Repita as etapas (3) e (4) várias vezes para criar uma distribuição aleatória das diferenças nas proporções simuladas.

Interpretando os resultados da simulação

Simulando o experimento sob o pressuposto de independência, ou seja, deixando as coisas ao acaso.

Se os resultados das simulações baseadas no modelo nulo se parecem com os dados, então podemos determinar que a diferença entre as proporções das estimativas corretas nos dois grupos foi simplesmente *devido ao acaso*.

Se os resultados das simulações baseadas no modelo nulo não se parecem com os dados, então podemos determinar que a diferença entre as proporções dos palpites corretos nos dois grupos não foi devida ao acaso, mas *sim porque as pessoas realmente conhecem o dorso de suas mãos melhor do que a palma*.

Resultados simulados

- No próximo slide você pode ver o resultado de um teste de hipótese (usando apenas 100 simulações para manter os resultados simples).
- Cada ponto representa uma diferença na proporção simulada de sucessos. Podemos ver que a distribuição está centrada em 0 (o valor nulo).
- Também podemos ver que 9 das 100 simulações produziram diferenças simuladas pelo menos tão grandes quanto a diferença observada (p-valor = 0.09).

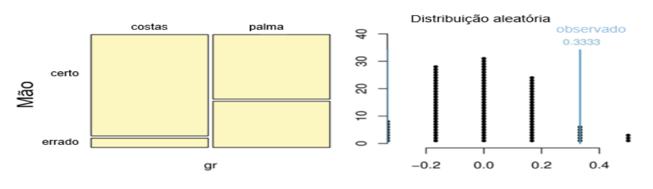
Simulação

```
mão = as.factor (c (rep ("correto", 7), rep ("errado", 5), c (rep ("correto", 11), rep ("errado", 1))))
gr = c(rep("palma",12),rep("dorso",12))
inferência (mão, gr, est = "proporção", tipo = "ht", null = 0, alternativa = "twosided",
        order = c ("voltar", "palma"), sucesso = "correto", método = "simulação", semente = 879,
        nsim = 100)
```

Simulação

```
mão = as.factor (c (rep ("correto", 7), rep ("errado", 5), c (rep ("correto", 11), rep ("errado", 1))))
gr = c(rep("palma",12),rep("dorso",12))
inferência (mão, gr, est = "proporção", tipo = "ht", null = 0, alternativa = "twosided",
        order = c ("voltar", "palma"), sucesso = "correto", método = "simulação", semente = 879,
        nsim = 100)
```

```
Variável de resposta: categórica, Variável explicativa: categórica
Diferença entre duas proporções - sucesso: correto
Estatísticas resumidas:
         x
y          soma  palma  dorso
  correto   11     7  18
  errado      1     5   6
  soma       12    12  24
Diferença observada entre proporções (palma da mão) = 0.3333
H0: dorso - p_palma = 0
HA: p_dorso - p_palma != 0
valor-p =  0.18
```



Conclusão

Os resultados da simulação sugerem que as pessoas conhecem melhor o dorso das mãos do que a palma?
(Lembre-se: havia 33.3 % mais acertos no grupo ddo dorso nos dados observados.)

(a) Sim

(b) Não

p-valor = 0.09 $>$ 0.05, não rejeitar H_0. Os dados não fornecem evidências convincentes de que as pessoas conhecem melhor o dorso das mãos do que a palma das mãos.

Conclusão

Os resultados da simulação sugerem que as pessoas conhecem melhor o dorso das mãos do que a palma?
(Lembre-se: havia 33.3 % mais acertos no grupo ddo dorso nos dados observados.)

(a) Sim

(b) *Não*

p-valor = 0.09 $>$ 0.05, não rejeitar H_0. Os dados não fornecem evidências convincentes de que as pessoas conhecem melhor o dorso das mãos do que a palma das mãos.