

シンプソンのパラドックス

問題①

- ある病気を予防するのに効果が期待される 2 種類の薬A,薬Bとの効果の違いを比較したいとします。無作為化実験ができると良いのですが、観察データ（薬の効果分析.xlsx file）しか手に入らなかったとします。このデータを元に薬A,薬Bの効果に関する分析を行ってください。

分析① 性別別の効果

男性

	効果あり	効果なし	総計
A	81	6	87
B	234	36	270

女性

	効果あり	効果なし	総計
A	192	71	263
B	55	25	80

分析① 性別別の効果

男性

	効果あり	効果なし	総計
A	81	6	87
B	234	36	270

$$\frac{81}{87} = 93.1\%$$

女性

	効果あり	効果なし	総計
A	192	71	263
B	55	25	80

分析① 性別別の効果

男性

	効果あり	効果なし	総計
A	81	6	87
B	234	36	270

$$\frac{234}{270} = 86.7\%$$

女性

	効果あり	効果なし	総計
A	192	71	263
B	55	25	80

分析① 性別別の効果

男性

	効果あり	効果なし	総計
A	81	6	87
B	234	36	270

女性

	効果あり	効果なし	総計
A	192	71	263
B	55	25	80

$$\frac{192}{263} = 73\%$$

分析① 性別別の効果

男性

	効果あり	効果なし	総計
A	81	6	87
B	234	36	270

女性

	効果あり	効果なし	総計
A	192	71	263
B	55	25	80

$$\frac{55}{80} = 68.8\%$$

分析① 性別別の効果

男性

	効果あり	効果なし	総計
A	81(93.1%)	6(6.9%)	87
B	234(86.7%)	36(13.3%)	270

女性

	効果あり	効果なし	総計
A	192(73%)	71(27%)	263
B	55(68.8%)	25(31.3%)	80

男性・女性で薬Aの
効果があると結論

分析② 性別別に集計しない場合

	効果あり	効果なし	総計
A	273	77	350
B	289	61	350

分析② 性別別に集計しない場合

	効果あり	効果なし	総計
A	273	77	350
B	289	61	350

$$\frac{273}{350} = 78\%$$

分析② 性別別に集計しない場合

	効果あり	効果なし	総計
A	273	77	350
B	289	61	350

$$\frac{289}{350} = 82.6\%$$

分析② 性別別に集計しない場合

	効果あり	効果なし	総計
A	273(78%)	77(22%)	350
B	289(82.6%)	61(17.4%)	350

薬Bの効果があると結論

分析② 性別別に集計しない場合

	効果あり	効果なし	総計
A	273(78%)	77(22%)	350
B	289(82.6%)	61(17.4%)	350

薬Bの効果があると結論

集団全体を見た時とその小集団に注目した時で一見矛盾したような結論がデータから導かれてしまうことを指して、「シンプソンのパラドックス」と呼びます。

なぜ？

男性

	効果あり	効果なし	総計
薬1	A		B
薬2	a		b

$$\frac{A}{B} > \frac{a}{b}$$

女性

	効果あり	効果なし	総計
薬1	C		D
薬2	c		d

$$\frac{C}{D} > \frac{c}{d}$$

なぜ？

男性

	効果あり	効果なし	総計
薬1	A		B
薬2	a		b

$$\frac{A}{B} > \frac{a}{b}$$

	効果あり	効果なし	総計
薬1	A+C		B+D
薬2	a+c		b+d

女性

	効果あり	効果なし	総計
薬1	C		D
薬2	c		d

$$\frac{C}{D} > \frac{c}{d}$$

なぜ？

男性

	効果あり	効果なし	総計
薬1	A		B
薬2	a		b

$$\frac{A}{B} > \frac{a}{b}$$

	効果あり	効果なし	総計
薬1	A+C		B+D
薬2	a+c		b+d

女性

	効果あり	効果なし	総計
薬1	C		D
薬2	c		d

$$\frac{C}{D} > \frac{c}{d}$$

なぜ？

男性

	効果あり	効果なし	総計
薬1	A		B
薬2	a		b

$$\frac{A}{B} > \frac{a}{b}$$

女性

	効果あり	効果なし	総計
薬1	C		D
薬2	c		d

$$\frac{C}{D} > \frac{c}{d}$$

	効果あり	効果なし	総計
薬1	A+C		B+D
薬2	a+c		b+d

~~$$\frac{A+C}{B+D} > \frac{a+c}{b+d}$$~~

なぜ？

データから得られる結果

- ・「男性では、A薬のほうが効果があつた人の割合が高い」
- ・「女性では、A薬の治療のほうが効果があつた人の割合が高い」
- ・「集団全体では、B薬のほうが効果があつた人の割合が高い」

なぜ？

データから得られる結果

- ・「男性では、A薬のほうが効果があつた人の割合が高い」
- ・「女性では、A薬の治療のほうが効果があつた人の割合が高い」
- ・「集団全体では、B薬のほうが効果があつた人の割合が高い」

矛盾はない

「割合が高い」という表現は完全に数学的な表現であり、集団全体で関連の方向性が逆転したこと自体も数学的に起こりうる自然な現象。
したがって、この三つの結果は矛盾していないと考えることができる。

なぜ？

データから得られる結果を次のように解釈する

- ・「男性では、A薬のほうが効果が良い」
- ・「女性では、A薬の治療のほうが効果が良い」
- ・「集団全体では、B薬のほうが効果が良い」

なぜ？

データから得られる結果を次のように解釈する

- ・「男性では、A薬のほうが効果が良い」
- ・「女性では、A薬の治療のほうが効果が良い」
- ・「集団全体では、B薬のほうが効果が良い」

間違いは

「効き目がいい」という表現に、治療の種類と効果の有無の間の因果関係が想定されている

なぜ？

データから得られる結果を次のように解釈する

- ・「男性では、A薬のほうが効果が良い」
- ・「女性では、A薬の治療のほうが効果が良い」
- ・「集団全体では、B薬のほうが効果が良い」

間違いは

男性でも女性でも従来治療より効かないのに男女合わせると従来治療より効き目がいいというのは論理的に間違い

問題②

- 1973年、米国でトップランクの大学の1つとして知られる、カリフォルニア州立大学バークレー校の大学院に数千人もの学生が受験しました、しかしどれだけの学生が合格し、不合格になったのでしょうか？
- UCBAmissionsデータセットを使って分析を行います。このデータは、バークレー校のに入学または合格を拒否された学生の性別と6つの学部の情報データです。

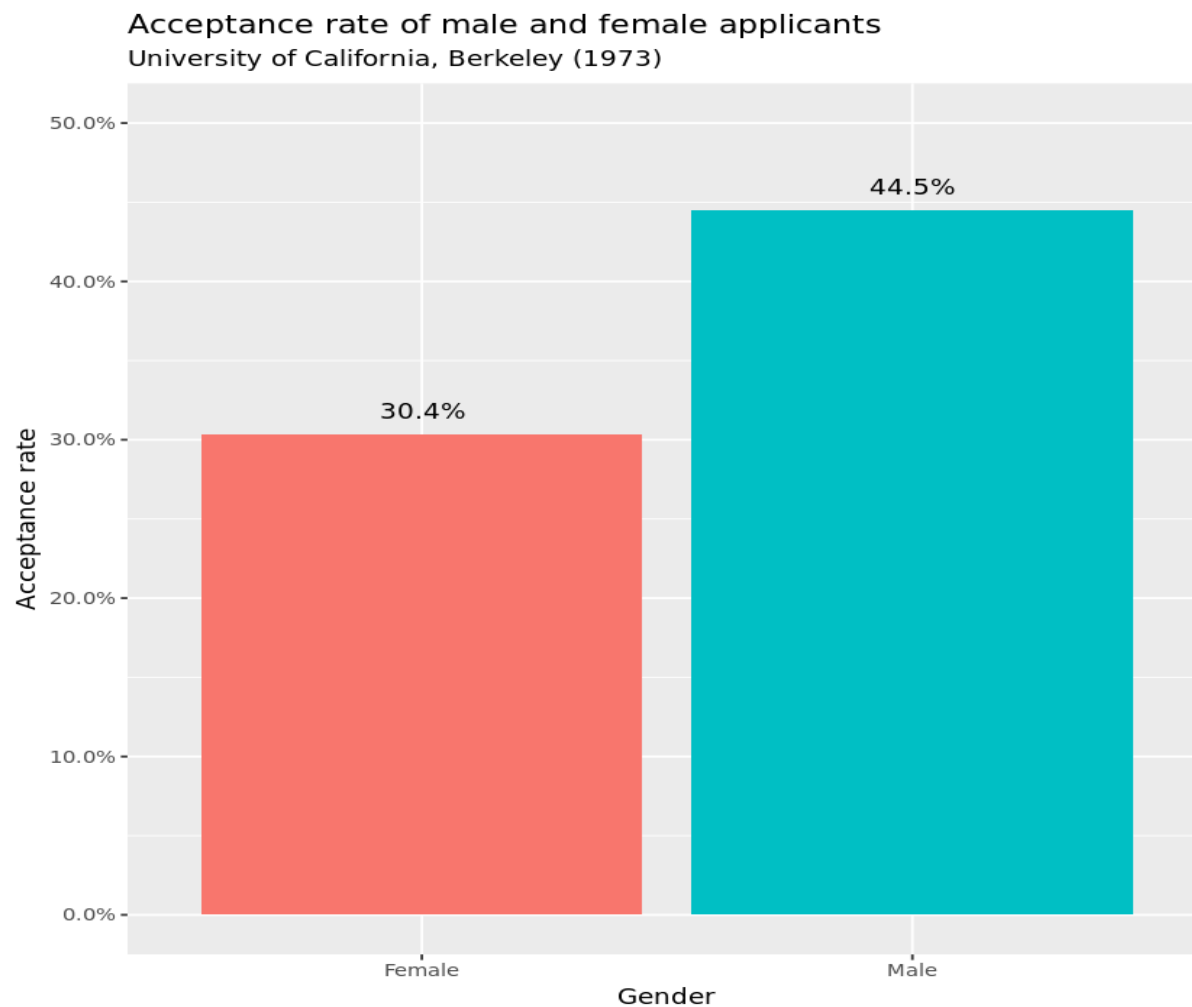
問題②

- 課題

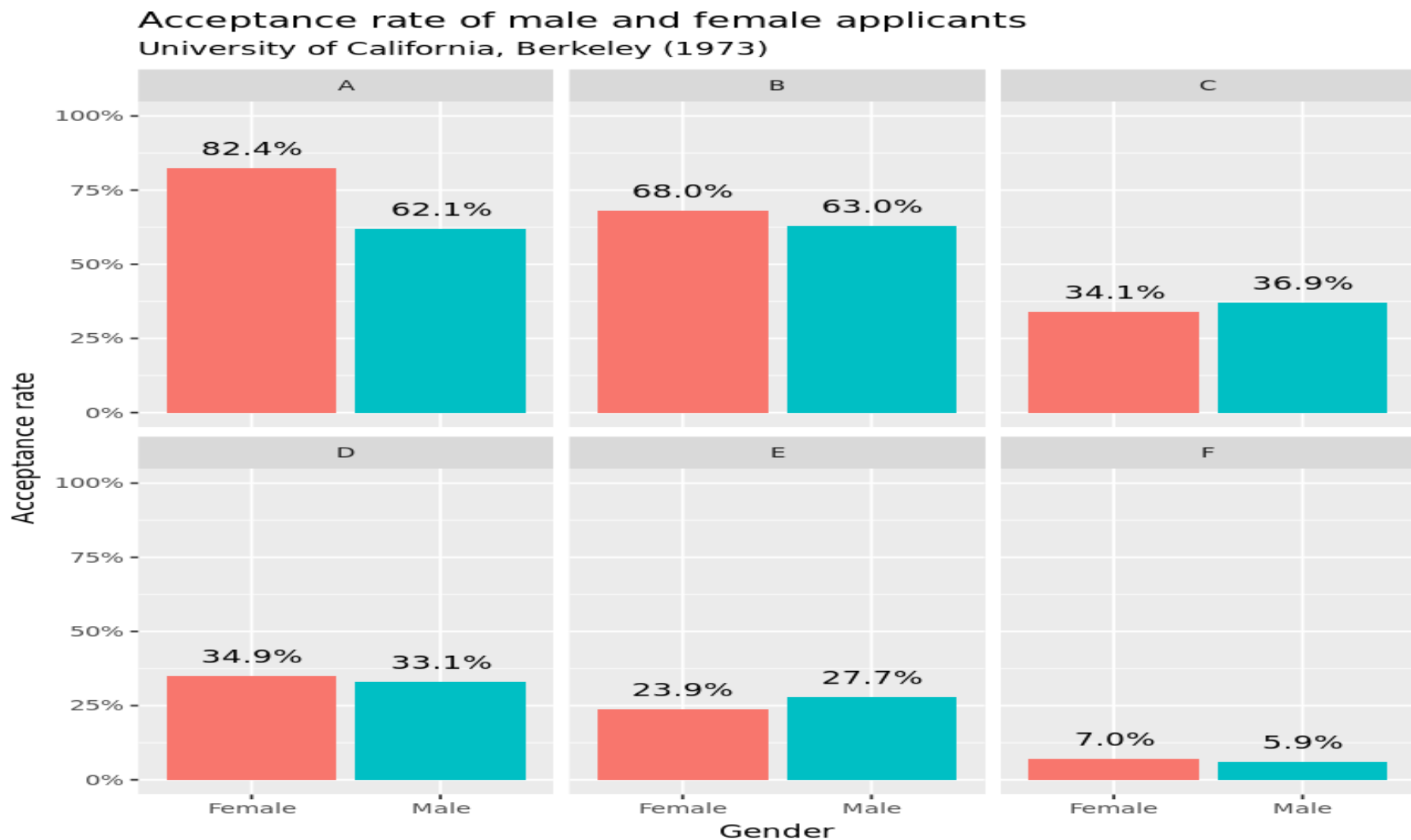
バークレー校の大学院の合否に性別が因子として関係があるかどうかを評価してください。

性別別合格率

	Admitted	Rejected	総計
Female	557	1278	1835
Male	1198	1493	2691
総計	1755	2771	4526

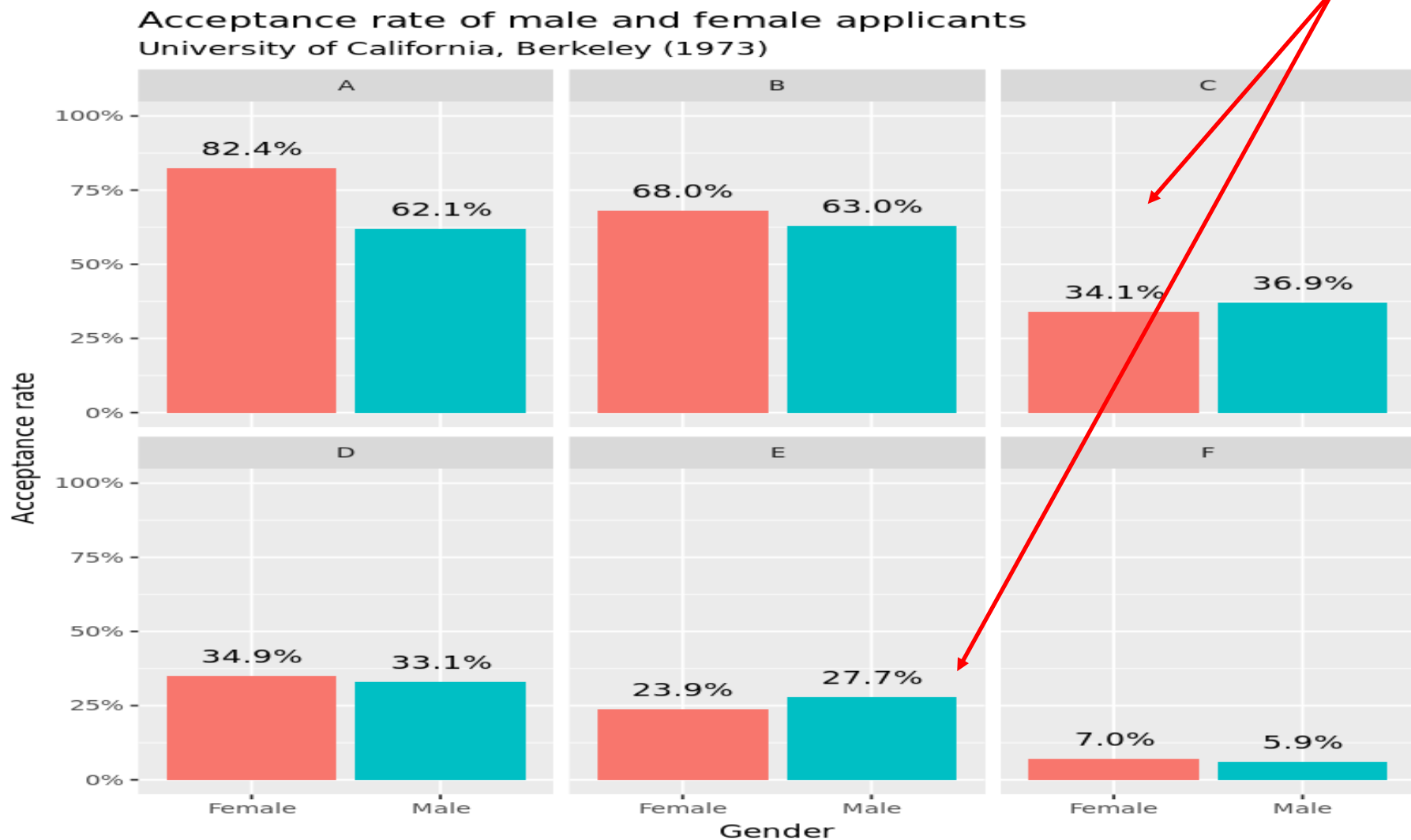


各学部における合格率



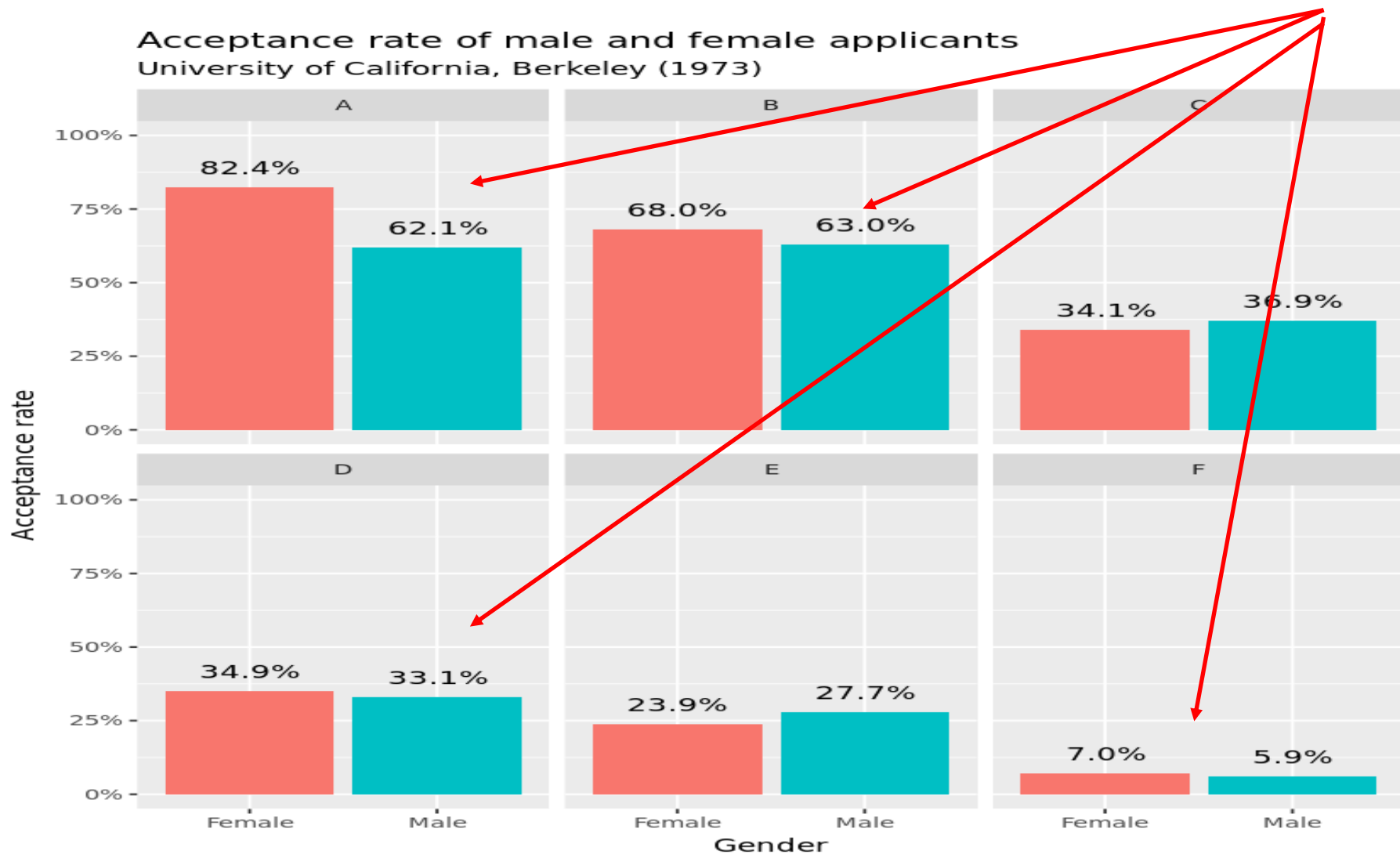
各学部における合格率

男性が合格しやすい

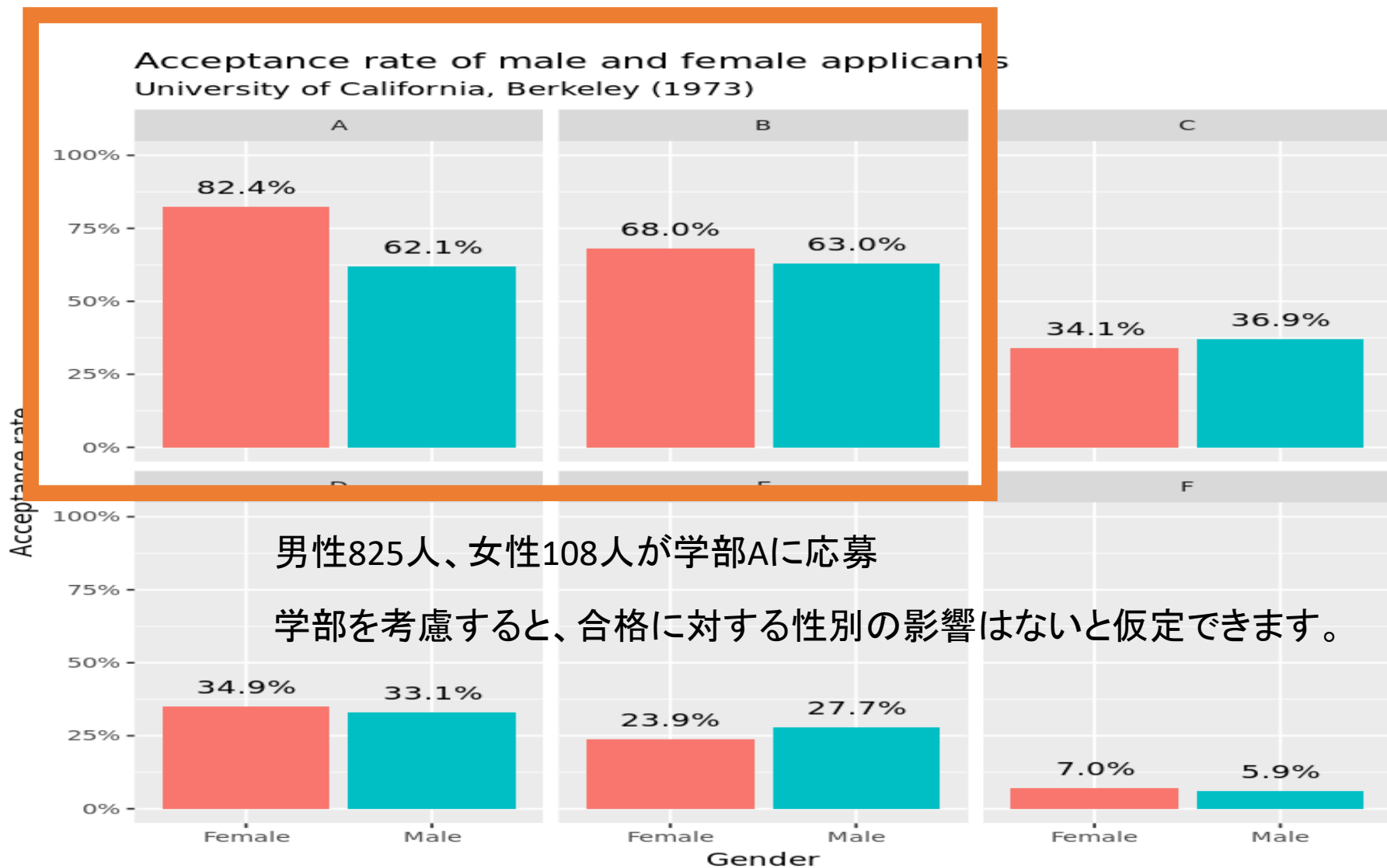


各学部における合格率

他の学部は女性の方が合格しやすい



各学部における合格率



ここまでで分かること

- 男性は学部CとEに入学する可能性が高い
- 女性是他のすべての部門に入学する可能性が高い。

最初の2つのプロットに注目すると

- 学部Aと学部Bが非常に簡単に合格できることがわかる。
- ただし、これらの学部に応募した女性は比較的少ない、
(男性825人、女性108人が学部Aに応募)

学部を考慮すると、合格に対する性別の影響はないと仮定できます。

ロジスティック回帰を使って検証

合否を性別だけで予測するロジスティック回帰モデル

```
Call:
glm(formula = 合否 ~ 性別, family = "binomial", data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5442  -1.2722   0.8506   1.0855   1.0855

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.83049    0.05077  16.357  <2e-16 ***
性別Male     -0.61035    0.06389  -9.553  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 6044.3  on 4525  degrees of freedom
Residual deviance: 5950.9  on 4524  degrees of freedom
AIC: 5954.9

Number of Fisher Scoring iterations: 4
```

男性の方が女性より $\exp(0.61035)=1.8$ 倍合格するオッズが高い

ロジスティック回帰を使って検証

合否を性別と学部で予測するロジスティック回帰モデル

```
Call:
glm(formula = 合否 ~ 性別 + 学部, family = "binomial",
    data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3613  -0.9588   0.3741   0.9306   1.4773

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.68192    0.09911  -6.880 5.97e-12 ***
性別Male      0.09987    0.08085   1.235  0.217
学部B         0.04340    0.10984   0.395  0.693
学部C         1.26260    0.10663  11.841 < 2e-16 ***
学部D         1.29461    0.10582  12.234 < 2e-16 ***
学部E         1.73931    0.12611  13.792 < 2e-16 ***
学部F         3.30648    0.16998  19.452 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 6044.3  on 4525  degrees of freedom
Residual deviance: 5187.5  on 4519  degrees of freedom
AIC: 5201.5

Number of Fisher Scoring iterations: 5
```

男性の方が女性より $\exp(-0.09987) = 0.9049551$ 倍合格するオッズが低い

シンプソンのパラドックス

- 最後に、シンプソンのパラドックスが機能していることがわかります。入学に対する学部の影響を制御すると、性別の影響はなくなります。
- 学部を管理すると、女性の合格確率が男性より高くなる可能性を示唆しています。ただし、この効果は統計的に有意ではないため ($p > 0.05$)、1973年にはどちらの性別の申請者に対してもキャンパス全体のバイアスはなかったと結論付けます。
- とはいえ、個々の部門が独自の入学プロセスを処理することが多いため、ある学部にはバイアスが存在し、別の学部にはバイアスが存在しない可能性があります。女性の82.4%が入学し、男性は62.1%しか入学できなかった学部Aを調べてみましょう。

学部Aの合格率

```
Call:
glm(formula = 合否 ~ 性別, family = "binomial", data = data2)
```

```
Deviance Residuals:
```

Min	1Q	Median	3Q	Max
-0.9768	-0.9768	-0.9768	1.3922	1.8642

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.5442	0.2527	-6.110	9.94e-10 ***
性別Male	1.0521	0.2627	4.005	6.21e-05 ***

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 1214.7  on 932  degrees of freedom
Residual deviance: 1195.7  on 931  degrees of freedom
AIC: 1199.7
```

```
Number of Fisher Scoring iterations: 4
```

学部Aは女性の方が入学しやすい

学部Aの合格率(カイ二乗検定)

帰無仮説「男女で合格率に差がない」
対立仮説「男女で合格率に差がある」

観測度数

	Admitted	Rejected
Female	89	19
Male	512	313

期待度数

	Admitted	Rejected
Female	69.569131	38.43087
Male	531.43086	293.5691

P値 : 3.2804E-05

結論 : 対立仮説「男女で合格率に差がある」

————→ 学部Aでは女子の合格率が男性より高い

学部Aでは男性を差別した？

- 女性の受験者の方が男性よりも優秀だった可能性もある
- この問題を扱ったHammel & O'Connell (1975) は差別を“受験者の能力でなく、性別によって合否が決定される”ことと定義した。
- 受験者の能力に関するデータが手元にないために、合格率の佐差が性差別による物だとは判断できない。