生体情報論演習
- 統計法の実践 第4回-

2011. 7. 1.
京都大学 情報学研究科 杉山麿人
仮説検定の例 1

• 自称超能力者がいた。彼は本物か？
  - 封筒の中の絵を当ててる。20 個中何個当てることができるか？
仮説検定の例 1

・1 個しか当てられなかったら超能力はうそ？
  2 個当てたら？3 個当てたら？20 個全部当てたら？
  → 超能力者かどうかを結論するには、どこかで基準を決めなければならない

・似た状況が、生命科学で現れる
  → コントロール実験において、コントロール群と処理群に差はあるのか？
仮説検定の例 2

<table>
<thead>
<tr>
<th>個体</th>
<th>コントロール群</th>
<th>処理群</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>39</td>
<td>83</td>
</tr>
<tr>
<td>2</td>
<td>26</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>69</td>
</tr>
<tr>
<td>4</td>
<td>56</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>55</td>
<td>59</td>
</tr>
<tr>
<td>6</td>
<td>43</td>
<td>51</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>70</td>
</tr>
<tr>
<td>8</td>
<td>57</td>
<td>62</td>
</tr>
<tr>
<td>9</td>
<td>45</td>
<td>67</td>
</tr>
<tr>
<td>10</td>
<td>36</td>
<td>84</td>
</tr>
</tbody>
</table>

mean ± s.e.m.
仮説検定の例 2

- 仮説：「コントロール群と処理群に差がある」は正しいか？
- 科学論文では、仮説が妥当かどうか結論する必要がある
- このとき、仮説検定を用いる
仮説検定の手順

・仮説が○の領域と×の領域の境界値（しきい値）を決める（たいてい0.01か0.05）
・コントロール群の分布から境界値を決める
・処理群の平均値が境界値より大きければ仮説は○, そうでなければ仮説は×

・実際にはExcelなどでP値を求める
・$P < 0.01(0.05)$なら差があると結論する
仮説検定の手順

この面積が0.01なるときに、○を示す。
境界値を超える場合は、×を示す。

反応

コントロール群
処理群
例題（1/3）

・目標: 血圧が上がる新薬Nを作った. 従来の薬Oより血圧が上がることを示したい

・10人の被験者に, それぞれNとOの2種類の薬を飲んでもらって, 効果を比較する

・2つの群を比較するときには, \( t \) 検定を用いる
  – 2つの群が独立なときには独立な\( t \) 検定
    • 群内, 群間の個体はすべて異なる
  – 2つの群が対応しているときには対応のある\( t \) 検定
    • 群内の個体は異なるが, 群間の個体は同じ
例題（2/3）

<table>
<thead>
<tr>
<th>被験者</th>
<th>前の薬O</th>
<th>新薬N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>105.9</td>
<td>114.9</td>
</tr>
<tr>
<td>2</td>
<td>81.8</td>
<td>143.7</td>
</tr>
<tr>
<td>3</td>
<td>86.5</td>
<td>157.5</td>
</tr>
<tr>
<td>4</td>
<td>92.1</td>
<td>137.5</td>
</tr>
<tr>
<td>5</td>
<td>107.3</td>
<td>154.9</td>
</tr>
<tr>
<td>6</td>
<td>107.1</td>
<td>156.0</td>
</tr>
<tr>
<td>7</td>
<td>106.1</td>
<td>158.3</td>
</tr>
<tr>
<td>8</td>
<td>116.2</td>
<td>109.6</td>
</tr>
<tr>
<td>9</td>
<td>101.8</td>
<td>113.8</td>
</tr>
<tr>
<td>10</td>
<td>98.6</td>
<td>135.6</td>
</tr>
</tbody>
</table>

mean ± s.e.m.
例題（3/3）

・データは対応している（同じ被験者が2種類の薬を飲んでいる）

・対応のあるt検定をおこなう

→ \( P = 0.000997 \)

・\( P < 0.01 \)なので、仮説は〇

結論: 新薬Nは従来の薬Oよりも効果がある
課題（1/2）

・データセット1, 2, 3それぞれについて、エラーバー付きの棒グラフ（mean ± s.e.m.）を並べて書け

・データセット1の平均値とデータセット2の平均値に差があるかどうかを、対応のあるt検定で比較せよ（両側）

・P値を示し、結論を述べる（基準は0.01）

・同様に、データセット1とデータセット3の平均値に差があるかどうかを比較せよ
課題（2/2）

• PowerPoint でまとめて提出
• 締め切り: 7月7日