

氏名

■ 正規分布（ガウス分布） データは嘘をつかない！

この話はドイツのハンブルクで本当に起こった話を元に行っている。（出典『数は魔術師』ジョージ・ガモフ+マーヴィン・スターン共著、由良統吉 訳、白揚社、1958 年、1999 年再版）

戦争に負けて品物が不足したドイツでは配給制度が行われ、パンの配給は 1 人 1 日に 200 g のパン 1 個だけとなってしまった。

毎日配給されるパンを受け取りにパン屋に行っている大学のゲーデ教授は、最近 30 日間に配給で受け取ったパンの重さを計ってきた。下の表はその重さの表である。

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 195 | 193 | 202 | 194 | 196 | 187 | 195 | 196 | 199 | 196 |
| 196 | 200 | 195 | 194 | 204 | 201 | 195 | 190 | 193 | 192 |
| 197 | 198 | 192 | 196 | 197 | 189 | 196 | 198 | 193 | 194 |

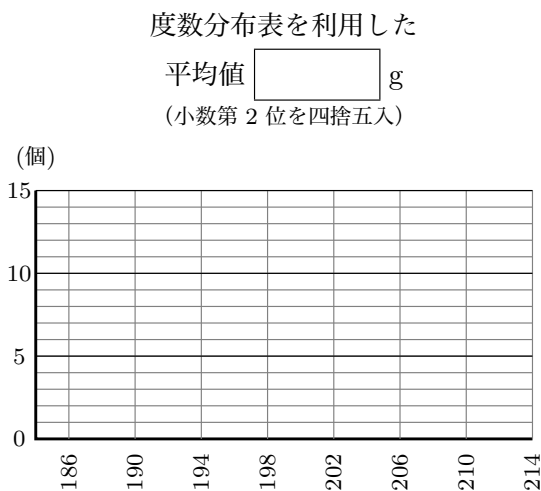
このデータを見たゲーデ教授はパン屋はわざと 200 g より小さいパンを作っていることを疑って、パン屋の主人に「君の店はわざと重さ不足のパンを作っているのではないか？。ちゃんと 200 g のパンを作るようにすべきだ。」と問いただした。

パン屋の主人は答えた。「すべてのパンをぴったり 200 g に作ることは無理です。多少のばらつきが出るため、200 g より軽いパンや重いパンができるのは仕方ありません。」と答えた。

しかしゲーデ教授は次のデータを示して反論した。

① 次の度数分布表とヒストグラムを作り平均値を求めて、ゲーデ教授が反論した理由を説明せよ。

| 階級 | 階級値 | 個数 | 階級値 × 個数 |
|-------------------|-----|----|----------|
| 186 g 以上～190 g 未満 | | | |
| 190 g 以上～194 g 未満 | | | |
| 194 g 以上～198 g 未満 | | | |
| 198 g 以上～202 g 未満 | | | |
| 202 g 以上～206 g 未満 | | | |
| 206 g 以上～210 g 未満 | | | |
| 210 g 以上～214 g 未満 | | | |
| 合計 | | 30 | |



ゲーデ教授のデータを示されたパン屋の主人はわざと 200 g より小さいパンを作っていることを認めて、今後は 200 g のパンを作ることを約束した。

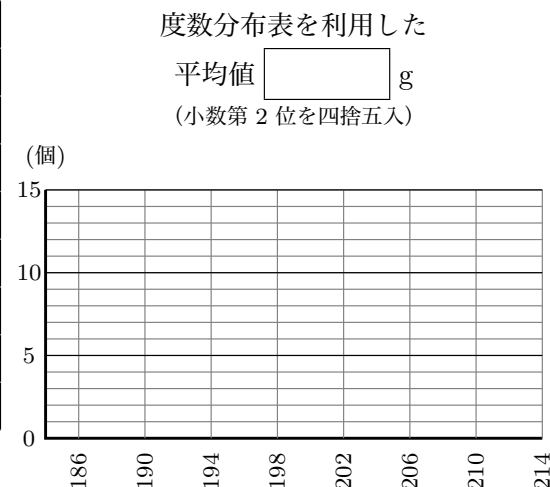
ゲーデ教授は、その後の 30 日間もパンの重さを計り続けた。下の表はその重さの表である。

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 201 | 202 | 202 | 205 | 202 | 207 | 201 | 200 | 200 | 204 |
| 206 | 200 | 201 | 201 | 204 | 200 | 202 | 204 | 200 | 205 |
| 200 | 201 | 201 | 210 | 201 | 206 | 201 | 200 | 202 | 204 |

このデータはパン屋はいまだに 200 g より小さいパンを作っていることを強く疑わせるものであった。

② 次の度数分布表とヒストグラムを作り平均値を求めて、ゲーデ教授がパン屋はいまだに 200 g より小さいパンを作っていると判断した理由を説明せよ。

| 階級 | 階級値 | 個数 | 階級値 × 個数 |
|-------------------|-----|----|----------|
| 186 g 以上～190 g 未満 | | | |
| 190 g 以上～194 g 未満 | | | |
| 194 g 以上～198 g 未満 | | | |
| 198 g 以上～202 g 未満 | | | |
| 202 g 以上～206 g 未満 | | | |
| 206 g 以上～210 g 未満 | | | |
| 210 g 以上～214 g 未満 | | | |
| 合計 | | 30 | |



氏名 _____

■ 検定 (統計学)

もし「10年に一度の大雨」が2年続けて起こったら、あなたは思うだろうか。確率 $\frac{1}{100}$ のことがたまたま起こったと思うだろうか。それとも「10年に一度の大雨」って嘘じゃないの? と思うだろうか。それを判定するのが**検定**である。

今回は **t 検定 (有意水準 5%の両側検定)** で判定してみよう。「200gのパンを作っている」という主張が正しいかどうかを「確率が5%未満のときは嘘である」「確率が95%以上のときは嘘とは言えない」と判定する。

データが30個なので、**自由度 29** (= 30 - 1) となって、**T 値**を計算して『2.045より大きいまたは -2.045より小さい』ときは「嘘である」と判定される。ExcelではTINV関数で計算できる。詳細は自分で調べてほしい。

■ 指摘前

(1) 30個のデータの平均値を求める

(2) 各データの偏差を求める

(3) 各データの偏差²を求める

(4) 偏差²の合計を求める

(5) 不偏分散 = $\frac{\text{イ}}{\text{データの個数} - 1} = \text{ウ}$

(6) 標準偏差 = $\sqrt{\text{ウ}} = \text{エ}$

(7) T 値 = $\frac{\text{ア} - 200 \text{ g}}{\left(\frac{\text{エ}}{\sqrt{\text{データの個数}}} \right)} = \text{オ}$

■ 指摘後

(1) 30個のデータの平均値を求める

(2) 各データの偏差を求める

(3) 各データの偏差²を求める

(4) 偏差²の合計を求める

(5) 不偏分散 = $\frac{\text{キ}}{\text{データの個数} - 1} = \text{ク}$

(6) 標準偏差 = $\sqrt{\text{ク}} = \text{ケ}$

(7) T 値 = $\frac{\text{カ} - 200 \text{ g}}{\left(\frac{\text{ケ}}{\sqrt{\text{データの個数}}} \right)} = \text{コ}$

前

| No | 重さ | 偏差 (重さ - 平均値) | 偏差 ² |
|----|-----|------------------|-----------------|
| 1 | 195 | | |
| 2 | 193 | | |
| 3 | 202 | | |
| 4 | 194 | | |
| 5 | 196 | | |
| 6 | 187 | | |
| 7 | 195 | | |
| 8 | 196 | | |
| 9 | 199 | | |
| 10 | 196 | | |
| 11 | 196 | | |
| 12 | 200 | | |
| 13 | 195 | | |
| 14 | 194 | | |
| 15 | 204 | | |
| 16 | 201 | | |
| 17 | 195 | | |
| 18 | 190 | | |
| 19 | 193 | | |
| 20 | 192 | | |
| 21 | 197 | | |
| 22 | 198 | | |
| 23 | 192 | | |
| 24 | 196 | | |
| 25 | 197 | | |
| 26 | 189 | | |
| 27 | 196 | | |
| 28 | 198 | | |
| 29 | 193 | | |
| 30 | 194 | | |
| 合計 | | | イ |

後

| No | 重さ | 偏差 (重さ - 平均値) | 偏差 ² |
|----|-----|------------------|-----------------|
| 1 | 201 | | |
| 2 | 202 | | |
| 3 | 202 | | |
| 4 | 205 | | |
| 5 | 202 | | |
| 6 | 207 | | |
| 7 | 201 | | |
| 8 | 200 | | |
| 9 | 200 | | |
| 10 | 204 | | |
| 11 | 206 | | |
| 12 | 200 | | |
| 13 | 201 | | |
| 14 | 201 | | |
| 15 | 204 | | |
| 16 | 200 | | |
| 17 | 202 | | |
| 18 | 204 | | |
| 19 | 200 | | |
| 20 | 205 | | |
| 21 | 200 | | |
| 22 | 201 | | |
| 23 | 201 | | |
| 24 | 210 | | |
| 25 | 201 | | |
| 26 | 206 | | |
| 27 | 201 | | |
| 28 | 200 | | |
| 29 | 202 | | |
| 30 | 204 | | |
| 合計 | | | キ |

② 階級値 188, 192, 196, 200, 204, 208, 212, 個数 0, 0, 0, 1, 15, 11, 3, 1, 度数利用平均値 202.7g

① データの分析# 4 階級値 188, 192, 196, 200, 204, 208, 212, 個数 2, 6, 15, 5, 2, 0, 0, 度数利用平均値 195.9g