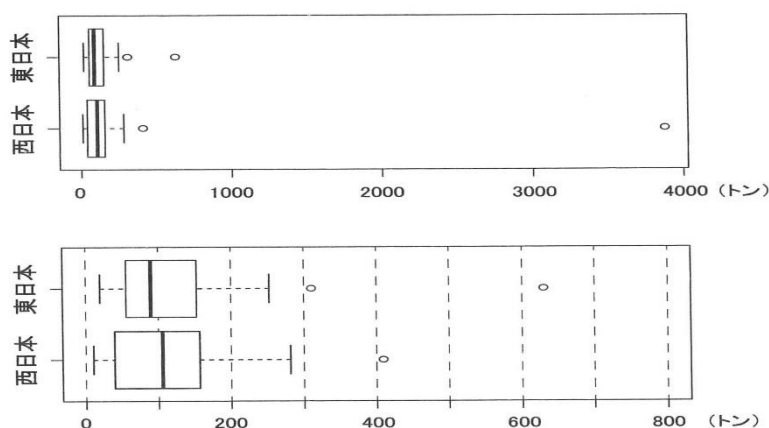


アメリカ式統計学 第六回 演習問題

※統計検定2級 2014年6月~2015年11月実施分より抜粋。

1. 箱ひげ図 (2015年6月実施)

次の箱ひげ図は、都道府県別のうどん（ゆでめん）の10万人当たりの生産量（トン）について、西日本（近畿以西の24府県）および東日本（23都道県）に分けて示したものである。なお2番目の図は、1番目の図を見やすくするために、横軸の値を0～800トンにして表示したものである。これらの箱ひげ図では、“「第3四分位数」+「四分位範囲」×1.5”以下の値をとるデータの最大値までひげを引き、それより大きな値を外れ値として○で示している。



資料：農林水産省「平成21年米麦加工食品生産動態等統計調査」および総務省統計局「平成22年国勢調査結果」

[1] 次の記述Ⅰ～Ⅲはこの箱ひげ図に関することである。

Ⅰ. 10万人当たりの生産量が最も少ない都道府県は西日本にある。

Ⅱ. 10万人当たりの生産量が最も多い都道府県は西日本にある。

Ⅲ. 10万人当たりの生産量が3番目に多い都道府県は西日本にある。

記述Ⅰ～Ⅲに関して、次の①～⑤のうちから最も適切なものを一つ選べ。

①Ⅰのみが正しい。 ②Ⅱのみが正しい。 ③ⅠとⅡのみが正しい。

④ⅠとⅢのみが正しい。 ⑤すべて正しい。

[2] この箱ひげ図がから読み取れることとして不適切なものを、次の①～⑤のうちから一つ選べ。

①600トン以上の都道府県数は、東日本と西日本で同数である。

②100トン以下の都道府県数は、西日本の方が多い。

③範囲は西日本の方が大きい。

④第1四分位数は東日本の方が大きい。

⑤西日本の府県の中に平均を100トン以上押し上げる外れ値があるので、平均は西日本の方が大きいと推察できる。

2. 統計量（2015 年 6 月実施）

国際通貨基金（IMF）によると、南米のベネズエラでは、2010 年の物価は 2009 年と比べて 1.282 倍、すなわち 2010 年の 1 年当たり物価上昇率は 28.2%であった。同様に、2011 年～2013 年の一年当たり物価上昇率がそれぞれ 26.1%、21.1%、40.6%であった。2010 年～2013 年の 4 年間の平均物価上昇率は何%か。次の①～⑤のうちから最も適切なものを 1 つ選べ。

- ① 28.2% ② 28.8% ③ 29.0% ④ 29.2% ⑤ 29.4%

3. 統計量（2015 年 6 月実施）

あるスーパーマーケットで平日に買い物に来た客の 1 回の購入金額を食料品フロアと衣料品フロアで調べてみた。次の表は購入金額（税抜）に関して要約したものである。

	食料品（円）	衣料品（円）
最小値	10	300
第 1 四分位数	700	2400
中央値	1400	3800
第 3 四分位数	2500	8000
最大値	30000	120000
平均	1900	5200
標準偏差	2500	7200
客数	11000	2500

[1] 購入金額の分布は食料品でも衣料品でもほぼ同じ形状をしていた。分布の形状に関して、次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 第 1 四分位数と第 3 四分位数の間に観測値のほとんどが含まれる、一様型に近い形をしている。
- ② 中央値を中心とした左右対称のベル型をしている。
- ③ 平均を中心とした左右対称のベル型をしている。
- ④ 右の裾が長い形状をしている。
- ⑤ 左の裾が長い形状をしている。

[2] 第 1 四分位数と第 3 四分位数を用いたばら付きの説明に関して、次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 第 1 四分位数と第 3 四分位数の間には観測値のうち約 4 分の 1 が含まれる。
- ② 第 1 四分位数と第 3 四分位数の間には観測値のうち約半分が含まれる。
- ③ 第 1 四分位数と第 3 四分位数の間には観測値のうち約 4 分の 3 が含まれる。
- ④ 第 1 四分位数と第 3 四分位数の間には観測値のうちほとんどが含まれる。
- ⑤ 第 1 四分位数より小さい観測値は全体の 4 分の 1、第 3 四分位数より大きい観測値は 4 分の 3 だが、その間に入る観測値の割合はわからない。

[3] 次の文章は食料品と衣料品の購入金額のばらつきについて述べたものである。(ア)～(オ)にあてはまるものとして、もっとも①～⑤のうちから最も適切なものを一つ選べ。

購入金額の標準偏差は食料品は 2500 円、衣料品は 7200 円だから、購入金額のばらつきは衣料品のほうが大きいといえることができる。しかし、食料品に比べて衣料品は単価の高いものが多く、平均も食料品と比べて約 2.7 倍であるから、衣料品の標準偏差が大きくなることは当然である。このような場合に観測値のばらつきを比較するには(ア)を用いればよい。(ア)の定義式は(イ)である。食料品の(ア)は(ウ)、衣料品の(ア)は(エ)だから、購入金額のばらつきが相対的に大きいのは(オ)であるが、大きな差ではない。

① (ア) 変動係数 (イ) $\frac{\text{平均}}{\text{標準偏差}}$ (ウ) 0.76 (エ) 0.72 (オ) 衣料品

② (ア) 変動係数 (イ) $\frac{\text{標準偏差}}{\text{平均}}$ (ウ) 1.32 (エ) 1.38 (オ) 衣料品

③ (ア) 変動係数 (イ) $\frac{\text{標準偏差}}{\text{平均}}$ (ウ) 1.32 (エ) 1.38 (オ) 食料品

④ (ア) 標準得点 (イ) $\frac{\text{平均}}{\text{標準偏差}}$ (ウ) 0.76 (エ) 0.72 (オ) 食料品

⑤ (ア) 標準得点 (イ) $\frac{\text{平均}}{\text{標準偏差}}$ (ウ) 1.32 (エ) 1.38 (オ) 食料品

4. クロス集計（2015 年 6 月実施）

次の表は、あるカレーライス専門店における客 400 人（男性 200 人、女性 200 人）を対象に実施したアンケートをまとめたクロス集計表である。この店では辛口と甘口の 2 種類のカレーライスを扱っている。

表 1: 全顧客

	満足	不満	合計
辛口を注文	116	84	200
甘口を注文	124	76	200
合計	240	160	400

表 2: 女性客のみ

	満足	不満	合計
辛口を注文	38	2	40
甘口を注文	122	38	160
合計	160	40	200

[1]男性客の中で辛口を注文して満足した客は何人いるか。次の①～⑤のうちから最も適切なものを一つ選べ。

①2 人 ②38 人 ③78 人 ④82 人 ⑤84 人

[2]この表からわかることの説明として、次の①～⑤のうちから最も適切なものを一つ選べ。

①全顧客の集計では、辛口でも甘口でも満足の割合は 60%程度であり変わらない。したがって、性別ごとに分けて集計する必要はなく、性別にかかわらず甘口と辛口には満足度の傾向にほとんど差がないことがわかる。

②甘口を注文した客のうち、満足の割合は女性客では 75%程度、男性客では 5%程度である。満足の割合がともに 0%でないということは、性別によらず辛口よりも甘口のほうが好まれていることを意味している。

③女性では、辛口の満足の割合は 95%だが、甘口は 75%である。男性客では、辛口の満足の割合は 50%程度、甘口は 5%である。性別によらず辛口の満足度が高いため、甘口よりも辛口を注文した人のほうが満足が高いといえる。

④女性で満足した人のうち甘口を注文したのは 75%程度なのに対し、男性で満足した人の中で甘口を注文した客はほとんどいなかった。したがって、甘口を注文する割合は性別によって大きく異なることがわかる。

⑤満足した客の割合は、女性客では 80%であったが男性客では 40%にとどまった。したがって、辛口と甘口に分けて集計する必要はなく、男性のほうが女性よりも不満であったことがわかる。

5. 幹葉図（2014 年 6 月実施）

次の図は、資産公開制度により 2007 年 10 月 27 日に公開された、内閣の政務官までを含む官僚 69 名の資産公開データの「家族分を含む総資産」を幹葉図として示したものである。

億 | 千万 (円)

0 | 0000011111 1112222233 3333333344 4444444555 5556667778 8899

1 | 011227

2 | 114

3 | 67

4 | 0

5 |

6 | 7

7 | 3

8 | 5

次の記述 I ～IIIはこの幹葉図に関するものである。

I.資産の中央値は 0.4 億円である。

II.資産の四分位範囲は 1.0 億円である。

III.資産の平均は 0.4 億円である。

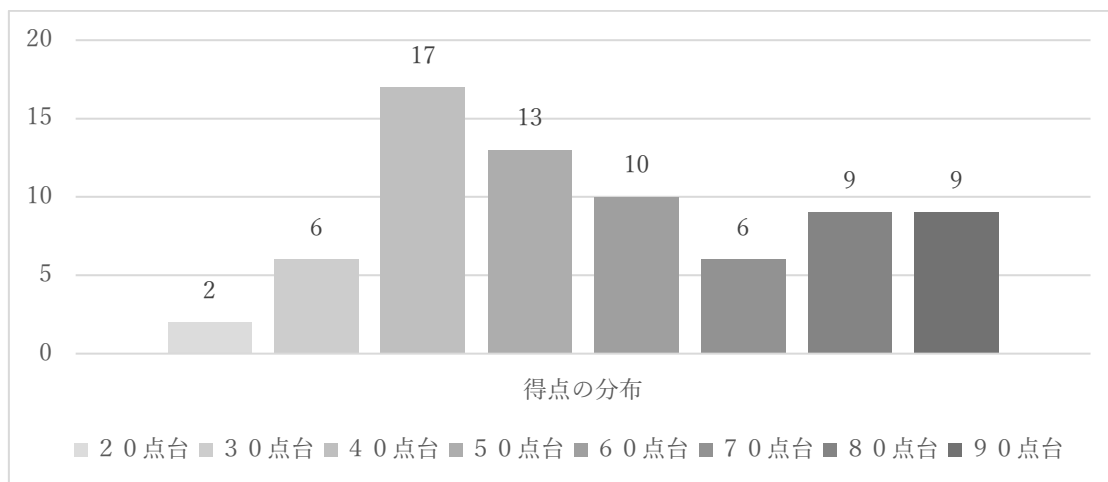
記述 I ～IIIに関して、次の①～⑤のうちから最も適切なものを一つ選べ。

① I のみ正しい。 ② II のみ正しい。 ③ III のみ正しい。

④ I と II のみ正しい。 ⑤ I と III のみ正しい。

6. ヒストグラム（2014 年 6 月実施）

次のヒストグラムは、あるテストを受けた 72 人の得点の分布を表したものである。20 点台の級（区間）は 20 点から 29 点の区間を、30 点台は 30 点から 39 点の区間を表し、その他も同様である。



[1]中央値はどの級にあるか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 40 点台 ② 50 点台 ③ 60 点台 ④ 70 点台 ⑤ 80 点台

[2]ヒストグラムをもとに算出した平均はおよそいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 40 点 ② 50 点 ③ 60 点 ④ 70 点 ⑤ 80 点

7. 確率（2014 年 11 月実施）

ある会社の待合室で、男性 8 人、女性 7 人の社員が会議のために部長を待ち、コーヒーか紅茶のどちらかを飲んでいて。男性でコーヒーを飲んでいてのは 5 人、女性でコーヒーを飲んでいてのは 2 人であった。

[1]部長が男性の 1 人にコーヒーか紅茶のどちらかを飲んでいてのか聞いたとき、その人が飲んでいてのがコーヒーであった確率はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ① $\frac{8}{3}$ ② $\frac{5}{8}$ ③ $\frac{2}{7}$ ④ $\frac{5}{7}$ ⑤ $\frac{3}{15}$

[2]部長が男性 1 人、女性 1 人に飲んでいてのものを聞いたときに、2 人が違う飲み物を飲んでいて確率はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ① $\frac{3}{7}$ ② $\frac{8}{15}$ ③ $\frac{25}{56}$ ④ $\frac{31}{56}$ ⑤ $\frac{31}{105}$

8. 【難】ポアソン分布（2014 年 11 月実施より一部抜粋）

次の表はある地域における 1 日の死亡者の集計結果（500 日間）である。

死亡者数(人)	0	1	2	3	4	5	6 以上	計
件数(日数)	55	144	140	95	45	15	6	500

[1] 死亡者数の平均値はいくらか。ただし、死亡者数が 7 人以上の日はなかった。次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 1.50 ② 1.75 ③ 2.00 ④ 2.25 ⑤ 2.50

[2] 1 日の死亡者数 X がパラメータ λ のポアソン分布に従うと仮定するとき、ある日の死亡者数が 3 人である確率を求める式として正しいものはどれか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ① λ ② $\lambda(1-\lambda)^3$ ③ ${}_6C_3 \lambda^3 (1-\lambda)^3$ ④ $\lambda e^{-3\lambda}$ ⑤ $\frac{\lambda^3 e^{-\lambda}}{3!}$

[3] 1 日の死亡者数 X がパラメータ λ のポアソン分布に従うと仮定するとき、 X^2 の期待値 $E(X^2)$ とパラメータ λ の関係として正しいものはどれか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ① $E(X^2) = \frac{2}{\lambda^2}$ ② $E(X^2) = \frac{1}{\lambda}$ ③ $E(X^2) = \lambda$ ④ $E(X^2) = \lambda^2 + \lambda$ ⑤ $E(X^2) = \lambda^2$

9. 期待値（2014 年 6 月出題）

袋の中に赤球 6 個、黄球 8 個、青球 a 個が入っている。この袋から球を 1 つ取り出したときに、それが赤球であるときは $X=1$ 、黄球であるときは $X=2$ 、青球であるときは $X=3$ とする。確率変数 X の期待値は 2 であるとするとき、次の問いに答えよ。

[1] a の値として、次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 3 ② 4 ③ 5 ④ 6 ⑤ 7

[2] X の分散はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 0.6 ② 0.7 ③ 0.8 ④ 0.9 ⑤ 1.0

10. 条件付確率（2014 年 6 月出題）

ある TV 局では、冬季スポーツのジャンプ競技とフィギュアスケート競技を放映した場合の視聴率を評価したいと考えている。過去のデータなどをもとにして、ジャンプ競技の視聴を事象 A、フィギュアスケート競技の視聴を事象 B とした場合に、各事象の確率は、 $\Pr(A)=0.3$ 、 $\Pr(B)=0.4$ 、 $\Pr(A \cup B)=0.5$ であることが分かっている。このとき、条件付確率 $\Pr(B|A)$ として、次の①～⑤のうちから最も適切なものを一つ選べ。

- ① 0.12 ② 0.15 ③ 0.3 ④ 0.5 ⑤ 0.67

11. ベイズ確率 (2014 年 6 月)

ある 100 円ショップでは、5 本入りのボールペンセットとして、A 社製と B 社製の 2 種類を調達している。全体の 60%を調達している A 社製のボールペンセットに不良品が含まれている確率は 0.01 であり、残りの 40%を調達している B 社製のボールペンセットに不良品が含まれている確率は、0.05 である。

購入したボールペンセットに不良品が含まれていた場合に、そのボールペンセットが B 社製のものである確率はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ①0.05 ②0.4 ③0.6 ④0.65 ⑤0.77

12. 期待値 (2015 年 11 月)

あたりを引く確率が $1/10$ のくじがある。引いたくじは毎回元に戻してから次のくじを引くものとする。このくじを、はじめて当たりが出るまで引き続ける。

[1]はじめて当たりが出るまでにくじを引く回数の期待値はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ①5 回 ②10 回 ③15 回 ④20 回 ⑤25 回

[2]5 回くじを引いてすべて外れであった。このとき、いまから当たりが出るまでにくじを引く回数の期待値はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ①5 回 ②10 回 ③15 回 ④20 回 ⑤25 回

13. 二項分布／正規分布 (2014 年 11 月出題)

1 から 8 までの目が等確率で出る正八面体のサイコロがある。このサイコロを 640 回振ったとき、1 の目が 90 回以上出る確率はいくらか。次の①～⑤のうちから最も適切なものを一つ選べ。

- ①0.09 ②0.12 ③0.15 ④0.18 ⑤0.21

14. 正規分布 (2014 年 6 月出題)

ある 100 点満点のテストの得点は、母集団全体で平均 60 点、標準偏差 15 点であった。A 君の得点は 75 点であり、B 君の得点は 70 であった。

[1]A 君の偏差値と B 君の偏差値に関する記述として、次の①～⑤のうちから最も適切なものを一つ選べ。

- ①A 君の偏差値は 75 であり、B 君の偏差値は 70 である。
②A 君の偏差値は 60 であり、B 君の偏差値は 56.7 である。
③A 君の偏差値は 63.3 であり、B 君の偏差値は 60 である。
④A 君の偏差値は 60 であり、B 君の偏差値は 50 である。
⑤A 君の偏差値は 60 であり、B 君の偏差値は 55 である。

[2]テストの得点は、平均 60 点、標準偏差 15 点の正規分布 $N(60, 15^2)$ に従うとする。500 人の受験者があったとき、A 君の得点と B 君の得点の間に入る受験者の人数はおよそ何人か。次の①～⑤のうちから最も適切なものを一つ選べ。

- ①50 人 ②100 人 ③150 人 ④200 人 ⑤250 人