

経済統計分析 [中間試験]

以下のすべてに答えなさい

多岐選択

(a) ~ (d) のうち、適切なものを 1 つ選びなさい。

1) $X = x$ を所与としたときの Y の条件付確率 $\Pr(Y = y|X = x)$ は

- (a) $\frac{\Pr(Y=y)}{\Pr(X=x)}$
- (b) $\sum_{i=1}^l \Pr(Y = y, X = x_i)$
- (c) $\frac{\Pr(X=x, Y=y)}{\Pr(Y=y)}$
- (d) $\frac{\Pr(X=x, Y=y)}{\Pr(X=x)}$

2) X と Y の相関係数は、...

- (a) 分散は常に正の値をとるから負の値をとりえない
- (b) 共分散の 2 乗で与えられる
- (c) X と Y の共分散を 2 つの標準偏差の積で割って計算される
- (d) $\frac{\text{cov}(X, Y)}{\text{var}(X)\text{var}(Y)}$ である。

3) 共分散がゼロに等しい多変量正規分布に従う変数については、...

- (a) 相関はほとんどの場合ゼロだが、そうなるとは限らない
- (b) 互いに独立である
- (c) 確率を計算するときには χ^2 分布を使うべきである
- (d) それぞれの変数の周辺分布は正規分布ではない

4) 自由度 ∞ の t 分布は...

- (a) 計算できない
- (b) 標準正規分布に等しい
- (c) 正規分布に似た釣鐘型の分布を持つが、裾が「重い」
- (d) 自由度 ∞ の χ^2 分布に等しい

5) 中心極限定理が述べているのは...

- (a) $\frac{\bar{Y} - \mu_Y}{\sigma_Y}$ の分布は標準正規分布によって近似される
- (b) $\bar{Y} \xrightarrow{p} \mu_Y$

- (c) \bar{Y} が $\mu_Y \pm c$ の範囲に入る確率が、任意の定数 $c > 0$ に対して、 n を大きくするといくらでも 1 にちかづく
- (d) だいたい $n > 30$ であれば、 t 分布は F 分布に収束する
- 6) 母平均 μ_Y の推定量 $\hat{\mu}_Y$ が一貫性を持つとは...
- (a) $\hat{\mu}_Y \xrightarrow{p} \mu_Y$
- (b) 誤差の 2 乗の平均が可能な限り小さい
- (c) Y が正規分布に従う
- (d) $\bar{Y} \xrightarrow{p} 0$
- 7) \bar{Y} の標準誤差 $SE(\bar{Y})$ は以下で与えられる
- (a) $\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$
- (b) s_Y^2/n
- (c) s_Y
- (d) s_Y/\sqrt{n}
- 8) Y_1, \dots, Y_n の加重平均で与えられるすべての不偏推定量のうち、 \bar{Y} は
- (a) μ_Y の唯一の一致推定量である
- (b) μ_Y の最も効率的な推定量である
- (c) 定義により分散を持たない数である
- (d) μ_Y の最も不偏な推定量である
- 9) t 統計量は...
- (a) $t = \frac{\bar{Y} - \mu_{Y,0}}{\sigma_Y^2/n}$
- (b) $t = \frac{\bar{Y} - \mu_{Y,0}}{SE(\bar{Y})}$
- (c) $t = \frac{(\bar{Y} - \mu_{Y,0})^2}{SE(\bar{Y})}$
- (d) 1.96
- 10) 相関係数は...
- (a) ゼロと 1 の間の値をとる
- (b) 線形関係をあらわす指標である
- (c) X が Y の原因になっているとき 1 に近い値をとる
- (d) 強い非線形関係をもっているとき大きな値をとる

記述問題 1

- 1) 確率変数 X と Y の相関係数は、各変数を線形変換しても変わらないことを示しなさい。
すなわち、線形変換された変数 $X^* = a + bX$ と $Y^* = c + dY$ (ただし、 a, b, c, d はゼロでない任意の定数) について、 $\text{corr}(X, Y) = \text{corr}(X^*, Y^*)$ を証明しなさい。
- 2) i.i.d. サンプルに対する次のような加重平均を母平均の推定量として考える。

$$\hat{Y} = \frac{1}{n} \left(\frac{1}{4}Y_1 + \frac{7}{4}Y_2 + \frac{1}{4}Y_3 + \frac{7}{4}Y_4 + \cdots + \frac{1}{4}Y_{n-1} + \frac{7}{4}Y_n \right)$$

この推定量 \hat{Y} が不偏推定量であることを示し、通常の標本平均 \bar{Y} に比べて有効でないことを示しなさい。

- 3) 2つの確率変数 X, Y に対して、任意の X の実現値に対して $E[Y|X] = 0$ が成り立つとき、 X と Y の相関係数がゼロになることを示しなさい。

記述問題 2

各都道府県の平均在院期間を調査したところ¹、秋田県の患者 128 人、広島県の患者 305 人についての無作為標本を得ました。

- 1) 秋田県の患者の在院日数の標本平均は 21.6 日、標本分散は 1256.7 (標準偏差 35.5) でした。在院日数の平均値は 20 日であるといってもよいか、検定しなさい。
- 2) 広島県の患者の在院日数の標本平均は 28.6 日、標本分散は 5643.7 (標準偏差 75.1) でした。在院日数の平均値は 30 日であるといってもよいか、検定しなさい。
- 3) 秋田県と広島県の患者の在院日数の平均値は等しいといってもよいか、検定しなさい。

¹2005 年の『患者調査』都道府県別の退院患者の在院日数データ (閲覧第 122 表) から数値をえましたが、あくまで仮想例です。