統計学Ⅱ　2021 年度秋学期　問題と解答

【注意】解答は解答用紙に記しなさい。必ず、学年、組、学籍番号、氏名を記入しなさい。＊印の問題は答えに至る途中過程も記しなさい(答えだけでは 0 点)。検定では仮説、検定統計量とその分布、判定、結論を明記しなさい。数字の解答は割り切れるものは小数点表示で、割り切れないものは分数で記しなさい。注意に従わない答案は減点します

【数表】

標準正規分布の右側 a%点 Z_a ($Pr(Z>Z_a)=a$)：$Z_5=1.645$ $Z_{2.5}=1.96$

自由度 m の χ^2 分布の右側 a%点 $W_a(m)$（$Pr(W>W_a)=a$）：$W_{97.5}(9)=3$、$W_{5.0}(9)=17$、$W_{2.5}(9)=20$

自由度 m,n の F 分布の右側 a%点 $F_a(m,n)$（$Pr(F>F_a)=a$）：$F_5(9,9)=3.18$、$F_5(10,10)=2.97$

自由度 m の t 分布の右側 a%点 $t_a(m)$（$Pr(t>t_a)=a$）：$t_5(9)=1.8$, $t_5(18)=1.7$, $t_{2.5}(9)=2.3$, $t_{2.5}(18)=2.1$

$\sqrt{2}≒1.4$　$\sqrt{6}≒2.5$

1 日吉駅の 20 件の家賃データを集めて家賃 Y(万円)を専有面積 X(m²)で説明するモデル $Y_i=\alpha+\beta X_i+\varepsilon_i$ を excel を用いて最小 2 乗法で推定し、推定式 $Y_i=a+bX_i+e_i$ について以下の結果を得た。(25)

Y_i ＝ 3.691＋0.1205X_i　　決定係数 0.1443　標準誤差 1.300

(2.992)　(1.742)　　(　　)内は t 値

(1)最小 2 乗法によって得られる望ましい推定量を何というか。漢字 9 字で書きなさい。

(2)上の式の標準誤差とは何に対する推定値か。

(3) t 値は何を検定する統計量かを説明しなさい。

(4)レポートでは 20 件の家賃順のデータを等間隔で幅広い範囲で集めたが、その利点を説明しなさい。

(5)「専有面積が広いと家賃が高くなる」という仮説を有意水準 5%で検定しなさい。

解答　各 5 点　正答率:1)36% 2)49% 3)47% 4)44% 5)77%

(1)最良線形不偏推定量　Best Linear Unbiased Estimator の訳語です

(2)誤差 ε_i の母標準偏差　残差 e_i の標準偏差は減点

(3)帰無仮説:係数＝0　係数の有意性も可

(4)係数の標準誤差を小さくするため　偏りをなくすためだと単なる無作為抽出でも良いはず

(5)H_0: $\beta=0$, H_1: $\beta>0$ とする。帰無仮説が真だとすると $t=b/s_b$ は自由度 18 の t 分布に従うが、観測値においては t=1.742>1.7 なので帰無仮説は有意水準 5%で棄却され、専有面積は有意に正の効果をもつといえる　仮説(H0：b=0 とか H1：β>1 など)や計算ミス等の誤りは適宜減点

2　第 2 回レポートでは N(50,100)の正規母集団から大きさ n の無作為標本($x_1,x_2,..,x_n$)の標本分散 s^2 の確率分布の乱数実験をおこなった。(25)

(1)標本分散 s^2 から作られる χ^2 分布に従う確率変数 W の算式を書きなさい。

(2*)この W の性質から s^2 の期待値 $E(s^2)$＝100 となることを示しなさい。

(3*)この W の性質から s^2 の分散 $Var(s^2)$の理論値を求める計算式を導出しなさい。

(4)n=51 とした時、この W は χ^2 分布以外のどのような確率分布に近似できるか。その平均と分散についても記すこと。

(5*)(4)の性質を利用して n=51 の時の s^2 が 130 を超える確率は 5%より大きいか、小さいかを答えよ。

解答　各 5 点　正答率:1)75% 2)39% 3)25% 4)45% 5)3%

(1)$W=(n-1)s^2/100$

(2)$E(W)=\frac{n-1}{100}E(s^2)=n-1$ より $E(s^2)=100$　E(W)=n-1 から展開されてないと 0 点

(3)$Var(W)=\{(n-1)/100\}^2Var(s^2)=2(n-1)$ より $Var(s^2)=$<u>20000/(n-1)</u>

(4)平均 50 分散 100 の正規分布　標準正規分布は 0 点

(5) $P(s^2>130)=P(W>\frac{50}{100}130=65)$。W～N(50,100)より $P(Z>(65-50)/10=1.5)>P(Z>1.645)=0.05$ <u>5%より大きい</u>

3　以下の問に答えなさい。(25)

(1*)X と Y がそれぞれ自由度 m, n の χ^2 分布に従うとき、X+Y が χ^2 分布に従うことを χ^2 分布の定義から説明しなさい。

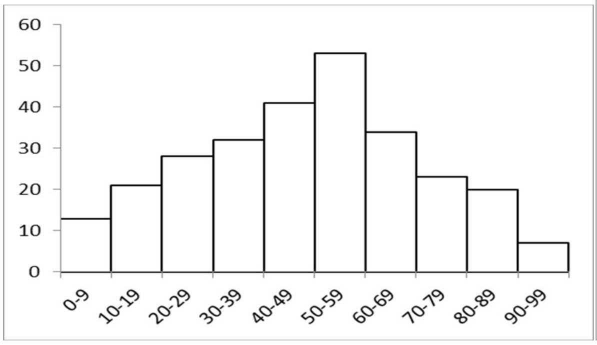
(2*)(1)の X+Y が χ^2 分布に従うことを積率母関数(MGF)から説明しなさい。ただし自由度 m の χ^2 分布の MGF は$(1-2t)^{-m/2}$とする。

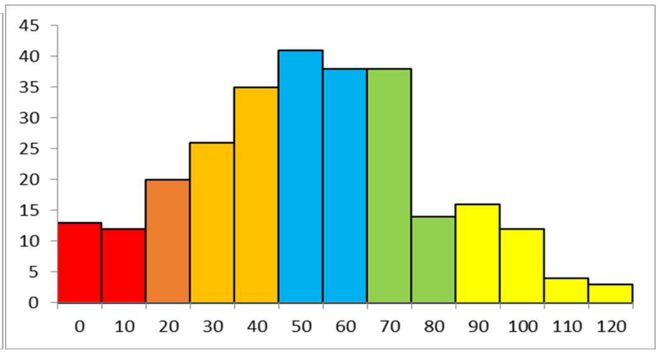
(3*)(1)の X の期待値 E(X)を MGF から導きなさい。
(4)ある保険の電話セールスの成約確率が 1%であるとき、ある支部で 1 日に 100 軒電話をかけたときの成約件数 x がポアソン分布に従う理由を説明しなさい。
(5*)この支部で 100 件電話して 2 件以上保険契約が成約する確率を求めよ。ただし、ポアソン分布の確率関数 f(x)=$e^{-\lambda}\frac{\lambda^x}{x!}$を利用し、e=3 として近似しなさい。
解答　各 5 点　正答率:1)19% 2)56% 3)41% 4)62% 5)55%
(1) z_i～N(0, 1)のとき X=$\sum_{i=1}^{m} z_i^2$, Y=$\sum_{i=1}^{n} z_i^2$だが、X+Y=$\sum_{i=1}^{m+n} z_i^2$なので X+Y は自由度 m+n の χ^2分布に従う
(2) $M_{x+y}(t)=M_x(t)M_y(t)=(1-2t)^{-(m+n)/2}$より X+Y は自由度 m+n の χ^2分布に従う
(3) E(x)=$M_x^{(1)}(t)\Big|_{t=0}=m(1-2t)^{-(m+2)/2}\Big|_{t=0}$=m
(4)成約件数 x は 2 項分布に従うが、n=100 と大きく、p=1%と小さいのでポアソン分布で近似できる　単位時間云々の説明は減点
(5) λ=1 のとき f(0)=1/3, f(1)=1/3 より P(x≧2)=1-f(0)-f(1)=1/3　簡単

4　近所のスーパー2 店(A 店,B 店)の LL サイズの卵をそれぞれ 1 パック(10 個)買って、卵 1 個あたり重量を測定したところ A 店、B 店それぞれの標本平均$\bar{x}$は 72.5g と 70g、標本標準偏差 s は$\sqrt{10}$g と$\sqrt{5}$g であった。(25)
(1*)この結果から A 店の卵 1 個あたり重量の母平均μ_Aの 95%信頼区間を求めよ。
(2*)この結果から A 店の卵 1 個あたり重量の母分散σ_A^2の 95%信頼区間を求めよ。
(3*)卵 1 個あたり重量の母分散は A 店＞B 店といえるか。有意水準 5%で検定を行いなさい。
(4*)卵 1 個あたり重量の母平均は A 店＞B 店といえるか。有意水準 5%で検定を行いなさい。
解答　(1)-(3)5 点(4)10 点　正答率:1)13% 2)69% 3)79% 4)64%
(1)t=$\frac{\bar{x}-\mu}{s/\sqrt{n}}$～t(n-1)より P[-2.3<t<2.3]=P[$\bar{x}-2.3\frac{\sqrt{10}}{\sqrt{10}}<\mu<\bar{x}+2.3\frac{\sqrt{10}}{\sqrt{10}}$]=P[$\underline{70.2}<\mu<\underline{74.8}$]=0.95　-1.96 は減点
(2)W=$(n-1)s^2/\sigma^2\sim\chi^2(9)$より Pr[3＜$(n-1)s^2/\sigma^2$＜20]=Pr[$(n-1)s^2/20<\sigma^2<(n-1)s^2/3$]=Pr[9・10/20＜$\sigma^2$＜9・10/3]=Pr[$\underline{4.5}<\sigma^2<\underline{30}$]=0.95　計算ミスは減点
(3)H_0: $\sigma_a^2=\sigma_b^2$　H_1: $\sigma_a^2>\sigma_b^2$　H_0が真だとすると F=s_a^2/s_b^2～F(9,9)に従うが、調査結果より F=10/5=2＜$F_5(9,9)$=3.18 なので有意水準 5%で H_0は採択され、散らばりは等しいといえる　F=5/10 は左片側検定で数表 F_{95}がないので誤り
(4)H_0: $\mu_a=\mu_b$ H_1: $\mu_a>\mu_b$　H_0が真だとすると t=$(\bar{x}_a-\bar{x}_b)/(s\sqrt{\frac{1}{n}+\frac{1}{n}})$は自由度 2(n－1)の t 分布に従うが、調査結果より $s^2=\frac{1}{2}10+\frac{1}{2}5$=15/2, $s\sqrt{\frac{1}{n}+\frac{1}{n}}=\sqrt{\frac{15}{2}\frac{1}{5}}=\frac{\sqrt{6}}{2}\fallingdotseq 1.25=\frac{2.5}{2}$。t=(72.5－70)/(2.5/2)=2＞ $t_5(18)$=1.7 なので H_0は棄却される。よって有意水準 5%で A 店の方が B 店より重いといえる　計算ミスは減点

総評
受験者数 272　平均 48.1　中央値 50　標準偏差 22.7　最低 0　最高 99
成績は、総点＝レポート乗数×期末得点+小テスト得点で、以下の基準で成績をつけます。
S：総点≧90　A：70≦総点＜90　B：50≦総点＜70　C：25≦総点＜50　D：総点＜25
（S と D の割合は学部指針に合わせるためにボーダーを決定しました。）
成績の人数分布は S：11%　A：17%　B：24%　C：24%　D：8%　未受験：16%　となりました。

期末得点分布(左は期末得点、右は総点)


0
10
20
30
40
50
60
0-9
10-19
20-29
30-39
40-49
50-59
60-69
70-79
80-89
90-99


0
5
10
15
20
25
30
35
40
45
0
10
20
30
40
50
60
70
80
90
100
110
120