2019.11.24

人文略解

人文１

[1] A 君の得点は 108 点で，全受験者の中で上位 34.46%，合格者の中で上位 68.92%，入学者の中で上位 61.15%に位置する。

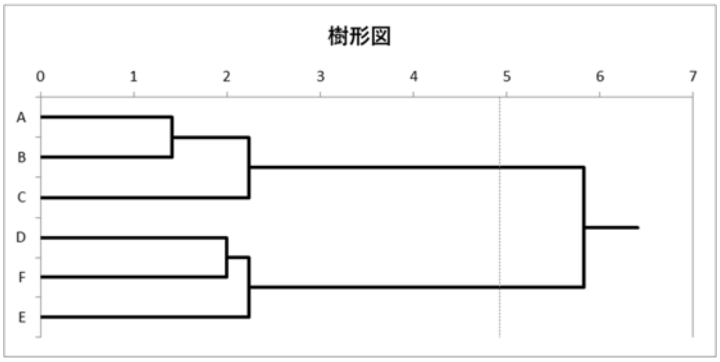
[2] 入学者の最低点は 100 点，最高点は 125.6 点である。

[3] $\dfrac{1}{10}\varphi\left(\dfrac{x-100}{20}\right) \quad (x \geq 100)$

[4] 合格者の得点の条件付き期待値は 115.96 であり，条件付き分散は 145.35 となる。

人文２

[1]



[2] クラスター形成距離 2.24～5.83 で分ける。

[3] 新たな代表点の座標は，クラスター1：(8, 9)，クラスター2：(3.5, 4) であり，新たなクラスターに属するマンション ID は，「クラスター1 : D，E，F」，「クラスター2 : A，B，C」となる．

[4] 初期値依存性とは，選択される初期代表点によって計算結果のクラスターが大きく異なる場合があることで，回避策として，異なる初期値を用いて何度か非階層的クラスタリングを実行し，結果が安定しているか調査することがあげられる。

人文３

[1]　$\theta = 0$ の参加者にとって正答確率が等しいのは項目 2 と項目 3 である。

[2]　項目 4 が最も正答しやすい項目である。

[3]　偏導関数は $\dfrac{\partial P_j(\theta)}{\partial \theta} = P_j(\theta)\{1 - P_j(\theta)\}a_j$ であり，$P_j(\theta) = 0.5$ のとき $\dfrac{a_j}{4}$ になる。

[4]　項目 j についての項目情報関数であり，当該項目による θ についての推定（測定）の精度を表す。

[5]　追加されたパラメータ c_j は当て推量パラメータと呼ばれ，項目 j に対して参加者が単に当て推量で解答したときに偶然正答できる確率を表す。

人文４

[1]　図に示されているのは標準解である。

[2]　f_1 から f_2 への直接効果の推定値は .640 で，間接効果は 0 であり，総合効果の推定値は .640 である。

[3]　f_1 から f_3 への直接効果の推定値は .014，間接効果の推定値は .020 である。したがって総合効果の推定値は $.014 + .020 = .034$ である。

[4]　$A = \begin{pmatrix} 0 & 0 & 0 \\ .640 & 0 & 0 \\ .014 & .032 & 0 \end{pmatrix}$，$\boldsymbol{e} = \begin{pmatrix} f_1 \\ e_{10} \\ e_{11} \end{pmatrix}$ である。

[5]　同値モデルとは，異なるモデルであっても同じ分散共分散行列を与える複数のモデルのことである。同値モデルの間では，適合度による統計的モデル選択を行うことができない。

2019.11.24

社会略解

社会１

[1] $E[\hat{Y}]=\bar{Y}$, $V[\hat{Y}]=\sum_{h=1}^{L} W_h{}^2 S_h{}^2\left(\frac{1}{n_h}-\frac{1}{N_h}\right)$

[2] $V[\hat{Y}_{(1)}]=\frac{1-f}{n}\sum_{h=1}^{L} W_h S_h{}^2$

[3] $\frac{n_h}{n}=\frac{W_h S_h}{\sum_{h=1}^{L} W_h S_h}=\frac{N_h S_h}{\sum_{h=1}^{L} N_h S_h}$

[4] $V[\hat{Y}_{(2)}]=\frac{1}{n}\left(\sum_{h=1}^{L} W_h S_h\right)^2-\frac{1}{N}\sum_{h=1}^{L} W_h S_h{}^2$

[5] 以下の 2 つの不等式から示される。

$$V[\hat{Y}_{(0)}]\approx V[\hat{Y}_{(1)}]+\frac{1}{n}\sum_{h=1}^{L} W_h(\bar{Y}_h-\bar{Y})^2\geq V[\hat{Y}_{(1)}]$$

$$n(V[\hat{Y}_{(1)}]-V[\hat{Y}_{(2)}])\approx \sum_{h=1}^{L} W_h(S_h-\bar{S})^2\geq 0$$

社会２

[1] ローレンツ曲線は所得分布の不平等度を曲線によって表している。ジニ係数はその不平等度を数値で表したものである。

[2] $G(u)=P(U\leq u)=u$ は区間 $(0, 1)$ 上の一様分布の累積分布関数である。

[3] ローレンツ曲線は $L(u)=1-(1-u)^{1-(1/\alpha)}$ $(0\leq u\leq 1)$ となり，ジニ係数は $G=\frac{1}{2\alpha-1}$ となる。

[4] Pareto(σ_1, α_1) が Pareto(σ_2, α_2) をローレンツ優越するための必要十分条件は

$\alpha_1>\alpha_2$ （σ_1，σ_2 は任意）

となる。

[5]　$u = F(x) = 1 - e^{-x/\mu} (x > 0)$ を得る。これは期待値 μ の指数分布の累積分布関数である。

社会３

[1]　$V[X_t] = \tau^2 = \dfrac{\sigma^2}{1-\phi^2}$，$Cov[X_t, X_{t+k}] = \dfrac{\sigma^2}{1-\phi^2}\phi^k$ であるので，$Cov[X_i, X_j] = \dfrac{\sigma^2}{1-\phi^2}\phi^{|i-j|}$ となる。自己相関係数は自己共分散を自己分散で割って $\rho_{ij} = \phi^{|i-j|}$ となる。

[2]　A は $\dfrac{1}{1-\phi^2}R$ すなわち $\dfrac{1}{\sigma^2}T$ の逆行列となり，$|A| = 1 - \phi^2$ となることが示される。また，$|R| = (1-\phi^2)^{n-1}$ である。

[3]　$Q_A = (1-\phi^2)x_1{}^2 + \sum_{i=2}^{n}(x_i - \phi x_{i-1})^2$ となり，すべての $\boldsymbol{x} \neq \boldsymbol{0}$ に対して $Q_A > 0$ となる。

[4]　$Q_B = \sum_{i=2}^{n}(x_i - \phi x_{i-1})^2$ となり，すべての ϕ に対し，$\boldsymbol{x} \neq \boldsymbol{0}$ であれば $Q_B \geq 0$ となる。$Q_B = 0$ となるのは，c を定数として $\boldsymbol{x} = c(1, \phi, \phi^2, ..., \phi^{n-1})'$ の場合である。

[5]　$s^2 = (1-\phi^2)x_1{}^2 + \sum_{i=2}^{n}(x_i - \phi x_{i-1})^2$ とし，$\chi_n(0.025)$ および $\chi_n(0.975)$ をそれぞれ自由度 n のカイ 2 乗分布の上側および下側 2.5%点として，σ^2 の 95%信頼区間は

$$\frac{s^2}{\chi_n^2(0.025)} \leq \sigma^2 \leq \frac{s^2}{\chi_n^2(0.975)}$$

と求められる。

社会４

[1]　$E[X]=\sum_{x=0}^{\infty}x\cdot\frac{\lambda^x}{x!}e^{-\lambda}=\lambda$ であり，$E[X(X-1)]=\lambda^2$ より $V[X]=\lambda$ を得る。

[2]　対数尤度関数 $l(\lambda)=n\overline{x}\log\lambda-n\lambda-\sum_{i=1}^{n}\log x_i!$ を λ で微分して 0 と置き

$$l'(\lambda)=\frac{n\overline{x}}{\lambda}-n=0$$

より最尤推定値 $\hat{\lambda}=\overline{x}$ を得る。

[3]　ポアソン分布のパラメータの推定値を $\hat{\lambda}$ とすると，サンプルサイズが n のとき，$x=k$ の期待度数は $e_k=nf(k;\hat{\lambda})=\frac{n\hat{\lambda}^k}{k!}e^{-\hat{\lambda}}$ となる。よって，$Y=\sum_{k=0}^{K}\frac{(f_k-e_k)^2}{e_k}$ とし，ポアソン性の帰無仮説の下で Y が近似的に自由度 $K-1$ のカイ 2 乗分布に従うことを利用した適合度のカイ 2 乗検定を行う。

[4]

(i)　対数尤度関数は,

$$l(\lambda,\omega)=m\log\omega+(n-m)\log(1-\omega)+A\log\lambda-(n-m)\lambda-\log\prod_{i=1}^{n-m}x_i!$$

で与えられる

(ii)　$\lambda=\frac{A}{n-m},\quad \omega=\frac{m}{n}$

(iii) λ の適当な初期値 $\lambda^{(0)}$ を選択し,

$$\begin{cases} m^{(t)}=\dfrac{f_0-n\exp[-\lambda^{(t-1)}]}{1-\exp[-\lambda^{(t-1)}]} \\ \lambda^{(t)}=\dfrac{A}{n-m^{(t)}} \end{cases}$$

なる反復計算式（EM アルゴリズム）が得られる。

[5]　$\tilde{\lambda}=1.98$ とすると，$\tilde{\omega}=0.0946$ と推定され，$x=0$ となった度数 $f_0=22$ での割合は $\frac{9.46}{22}=0.43$ となる。

2019.11.24

理工略解

理工１

[1]　部分積分を用いて $E[T]=\int_0^\infty tf(t)dt=-[tS(t)]_0^\infty+\int_0^\infty S(t)dt$ より導かれる。

[2]　部分積分により $\int_t^\infty S(x)dx=S(t)E[T-t\,|\,T>t]$ となることから与式を得る。また,

$$\int_0^\infty \exp[H(t)-H(t+x)]dx=\frac{1}{S(t)}\int_t^\infty S(y)dy=m(t)$$

となる。微分方程式 $\frac{S'(t)}{S(t)}=-\frac{1+m'(t)}{m(t)}$ の両辺を積分して $\log S(t)=-\int_0^t\frac{1+m'(x)}{m(x)}dx$ より与式が示される。

[3]

(1)　ハザード関数 $h(t)$ は，時刻 t まで稼働していた製品が時刻 t で故障する瞬間故障率を表す。

(2)　$H(t)$ が凸関数のとき，その導関数であるハザード関数 $h(t)$ は増加関数であることから T の確率分布は IFR である。$H(t)$ が凹関数の場合は，導関数は t の減少関数であることから，T の分布は DFR である。

[4]

(1)　確率密度関数は $g_\beta(t)=\beta t^{\beta-1}\exp[-t^\beta]$，ハザード関数は $h_\beta(t)=\beta t^{\beta-1}$ で与えられる。$\beta>1$ では分布は IFR，$\beta<1$ では DFR となる。$\beta=1$ ではハザード関数は定数で IFR かつ DFR となる。

(2)　$\beta=1/2$ とすると，確率密度関数とハザード関数はそれぞれ

$$g_{1/2}(t)=\frac{1}{2\sqrt{t}}\exp[-\sqrt{t}],\quad h_{1/2}(t)=\frac{1}{2}t^{-1/2}=\frac{1}{2\sqrt{t}}$$

であり，$\beta=2$ とすると，確率密度関数とハザード関数はそれぞれ

$$g_2(t)=2t\exp[-t^2],\quad h_2(t)=2t$$

となる。

※確率密度関数は参考として示した。

理工２

[1]　$\hat{\sigma}_W = 0.0241$

[2]　$\hat{\sigma}_B \approx 0.0423$，もしくは $\hat{\sigma}_B = 0.0416$。

[3]　対応策として，鋼材による特性 A の変動は偶然変動であるとみなし，$\bar{X}$ 管理図の管理限界線の設定に，群間変動も含めた $\hat{\sigma}_{Total}$ を用いる。

[4]　$\mathrm{UCL} = \mathrm{CL} + 3.0\sqrt{\mathrm{CL}} = 1.62 + 3.0\sqrt{1.62} = 5.43$

[5] ポアソン分布のパラメータがウェハ上で均一でない分布を仮定することが考えられる。

理工３

[1]　因子 A，B の主効果の F 値は，それぞれ 2.5，2.0 となる。また，順序 2 が適切である。

[2]　交互作用 A×B の F 値は 1.0 となる。

[3]　順序 4 が適切である。

[4]　焼成回数は，完全無作為化実験の場合には 8 回，分割実験の場合には 4 回となる。

[5]　因子 A の F 値は 2.5，因子 B，C，D の F 値はそれぞれ 4.0，1.0，1.0 となる。

理工４

[1]　$V[X_t]=\tau^2=\dfrac{\sigma^2}{1-\phi^2}$，$Cov[X_t,X_{t+k}]=\dfrac{\sigma^2}{1-\phi^2}\phi^k$　より　$Cov[X_i,X_j]=\dfrac{\sigma^2}{1-\phi^2}\phi^{|i-j|}$　となる。自己相関係数は自己共分散を自己分散で割って　$\rho_{ij}=\phi^{|i-j|}$　となる。

[2]　A　は　$\dfrac{1}{1-\phi^2}R$　すなわち　$\dfrac{1}{\sigma^2}T$　の逆行列となり，$|A|=1-\phi^2$　となることが示される。また，$|R|=(1-\phi^2)^{n-1}$　である。

[3]　$Q_A=(1-\phi^2)x_1^2+\sum_{i=2}^{n}(x_i-\phi x_{i-1})^2$　となり，すべての　$\boldsymbol{x}\neq\boldsymbol{0}$　に対して　$Q_A>0$　となる。

[4]　$Q_B=\sum_{i=2}^{n}(x_i-\phi x_{i-1})^2$　となり，すべての　ϕ　に対し，$\boldsymbol{x}\neq\boldsymbol{0}$　であれば　$Q_B\geq 0$　となる。$Q_B=0$　となるのは，c　を定数として　$\boldsymbol{x}=c(1,\phi,\phi^2,\ldots,\phi^{n-1})'$　の場合である。

[5]　$s^2=(1-\phi^2)x_1^2+\sum_{i=2}^{n}(x_i-\phi x_{i-1})^2$　とし，$\chi_n(0.025)$　および　$\chi_n(0.975)$　をそれぞれ自由度　n　のカイ 2 乗分布の上側および下側 2.5%点として，σ^2　の 95%信頼区間は

$$\frac{s^2}{\chi_n^2(0.025)}\leq\sigma^2\leq\frac{s^2}{\chi_n^2(0.975)}$$

と求められる。

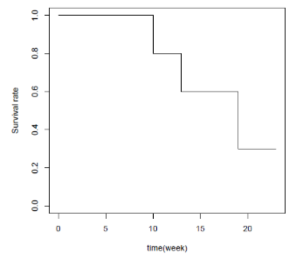
2019.11.24

医薬生物学略解

医薬１

[1]

生存時間（週）	n_j	d_j	$\hat{S}(t)$
10	5	1	0.8
13	4	1	0.6
19	2	1	0.3



[2]　次の関係式より求められる。

$$\exp\left(-\frac{d_k}{n_k}\right) \geq 1-\frac{d_k}{n_k} = \frac{n_k-d_k}{n_k}$$

[3]　T の確率密度関数を $f(t)$，累積分布関数を $F(t)$，生存関数を $S(t)$ としたとき，以下の関係式より導かれる。

$$E[X(\tau)] = E[\min(T,\tau)] = \int_0^\tau tf(t)dt + \int_\tau^\infty \tau f(t)dt = \int_0^\tau S(t)dt$$

[4]　分散は，$\mathrm{Var}[X(\tau)] = \dfrac{1-2\lambda\tau\exp(-\lambda\tau)-\exp(-2\lambda\tau)}{\lambda^2}$ となる。

[5]　ハザード λ の最尤推定値は

$$\hat{\lambda} = \frac{r}{\sum_{i=1}^n t_i} = \frac{3}{10+13+18+19+23} = \frac{3}{83} \approx 0.036$$

となる。また，$\tau = 20$ としたときの境界内平均生存時間と分散の推定値は

$$\hat{E}[X(\tau)] = \frac{1-\exp(-\hat{\lambda}\tau)}{\hat{\lambda}} \approx 14.239$$

$$\hat{V}[X(\tau)] = \frac{1-2\hat{\lambda}\tau\exp(-\hat{\lambda}\tau)-\exp(-2\hat{\lambda}\tau)}{\hat{\lambda}^2} \approx 48.018$$

である。

医薬２

[1]　リスク差の点推定値は 0.27 で，95%信頼区間は (0.23, 0.30) となる。

[2]　調整リスク差の推定値は 0.026 となる。

[3]　各層での傾向スコアの値は，層 1：0.143，層 2：0.550，層 3：0.550，層 4：0.853 となり，傾向スコアで条件付けた場合の共変量の分布は，$X=0$ と $X=1$ で等しいことが示される。

[4]　調整リスク差の推定値は 0.026 と上問 [2] で求めた値と一致する。

[5]　条件付き期待値の繰り返し公式を適切に応用することにより，与式が示される。

医薬３

[1]　検査法 A および B の感度は，それぞれ 0.8, 0.9 であり，陽性的中率はそれぞれ，0.8, 0.9 である。

[2]　疾患 D に罹患している群において，2 つの検査法で結果の異なる度数につき，二項分布 $B(23, 0.5)$ により検定すると，有意水準 0.05 で検定は有意ではなく，検査法の感度に差があるとはいえない。

[3]　$\sqrt{n}(\boldsymbol{p}-\boldsymbol{\pi})$ の分散共分散行列は $D-\boldsymbol{\pi}\boldsymbol{\pi}'$ で与えられる。ここで D は $\boldsymbol{\pi}$ の各成分を対角要素に持つ対角行列である。

[4]　$\sqrt{n}f(\boldsymbol{p})$ の分散は $\left(\dfrac{\partial f(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}\right)'(D-\boldsymbol{\pi}\boldsymbol{\pi}')\left(\dfrac{\partial f(\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}\right)$ となる。

[5]　$n=200$，$f(\boldsymbol{p})=-0.1$，$\hat{\sigma}^2=0.42$ より，$z=\dfrac{\sqrt{n}f(p)}{\sqrt{\hat{\sigma}^2}}=-2.18$ となって有意水準 5%で有意であり，検査法 A と検査法 B の陽性的中率の間には差があるといえる。

医薬４

[1] Student の t 検定の検定統計量は $t \approx 1.28$ となる。自由度 4 の t 分布の上側 2.5%点は 2.78 であるので，有意水準 5%の両側検定で有意ではない。検定の妥当性の条件は，1. 観測値の独立性，2. 分布の正規性，3. 等分散性，である。

[2] 薬剤 A 群の測定値の分布の分布関数を $F_1(x)$，薬剤 B 群の測定値の分布の分布関数を $F_2(x)$ としたとき，帰無仮説は $H_0 : F_1(\boldsymbol{x}) = F_2(x)$ である。

[3] 検定統計量は，両群のデータを込みにして順序を付けた場合の薬剤 A 群の順位和であり，$W_A = 13$ である（薬剤 B 群の順位和でもよい）。

[4] 右の表のようである。

順位和	確率
6	0.05
7	0.05
8	0.1
9	0.15
10	0.15
11	0.15
12	0.15
13	0.1
14	0.05
15	0.05

[5] 検定統計量の値は $W_A = 13$ であり，$P(13 \leq W_A \mid H_0) = 0.2$ であるので，両側 P 値は 0.4 となり，有意水準 5%で有意ではない。設問のデータで検定を行うと，すべての観測値で P 値は 0.05 未満にならず，検定は常に有意ではない。分布の離散性に起因する現象で，特にサンプルサイズが小さい場合には問題となる。

2019.11.24

共通略解

[1] 適合度のカイ 2 乗検定統計量の値は $Y = 4.0$ となる。自由度 3 のカイ 2 乗分布の上側 5%点は 7.81 であるので，有意水準 5%で帰無仮説は棄却されない。

[2] $k = 8$

[3] カテゴリー数が K の度数分布表の各度数 $f_1, \ldots, f_K$ は多項分布に従う。N が大きいとき，$f_1, \ldots, f_K$ の線形変換によって互いに独立に標準正規分布に従う確率変数を構成し，適合度のカイ 2 乗検定統計量 Y がそれらの 2 乗和となることを示して Y がカイ 2 乗分布に従うとする。

[4]

(i) 各比率の推定値は

$$\hat{r} = \frac{2f_{OO} + f_{AO} + f_{BO}}{2N}$$

$$\hat{p} = \frac{2f_{AA} + f_{AO} + f_{AB}}{2N}$$

$$\hat{q} = \frac{2f_{BB} + f_{BO} + f_{AB}}{2N}$$

となる（M ステップ）。

(ii) 観測されない度数の期待値は

$$E[f_{AA}] = Np^2, E[f_{AO}] = n_A - Np^2$$
$$E[f_{BB}] = Nq^2, E[f_{BO}] = n_B - Nq^2$$

で与えられる（E ステップ）。