人文 (18) 略解 (v-F.1)

人文１

[1] $f_n(x) = \dfrac{{}_6C_x \times {}_4C_{n-x}}{{}_{10}C_n}$ $(\max(0, n-4) \le x \le \min(n, 6))$。

$f_5(1) = \dfrac{6}{252} = 0.0238$, $f_5(2) = \dfrac{60}{252} = 0.2381$

[2]

(1) オッズ比：$(2 \times 1) / (3 \times 4)$＝1/6，ファイ係数：$\phi = \dfrac{2\times1-4\times3}{\sqrt{6\times4\times5\times5}} = -0.408$。両群間には違いがあるかもしれないが，観測数は少ないため値の信頼性には問題がある。

(2) 帰無仮説の下での期待度数の分割表は次のようになる。

群	中央値 7.5 より小さい	中央値 7.5 より大きい
A	3	3
B	2	2

(3) $Y \approx 1.667$，$Y' \approx 0.417$ となり，自由度 1 のカイ二乗分布の有意水準 10%の棄却限界値 2.71 以下であるので，有意水準 10%で帰無仮説は棄却できない。

(4) 片側 P 値は $0.0238 + 0.2381 = 0.2619$ となる。片側 P 値が既に 0.1 よりも大きいので，有意水準 10%で帰無仮説は棄却できない。

人文２

[1]　相関係数は 0.7 である。

[2]　境界線 U の方程式は $x_1 = 35$ である。生徒 a は進路 A，生徒 b と c は進路 B とアドバイスされる。

[3]　境界線 U と境界線 M の比較では，縦線が境界線 U で斜め線が境界線 M になる。この場合，生徒 a と b は進路 A，生徒 c は進路 B とアドバイスされる．

[4]　実際の誤判別の領域は最適誤判別の領域を含むため，実際の誤判別率は最適誤判別率より大きくなる。

人文３

[1]　X_2 が同じであれば，事前試験結果と補習試験結果が無相関であることを意味する。

[2]　X_2 と X_4 の間には X_1 による擬相関が，また X_3 と X_4 の間には X_2 による擬相関が生じている。

[3]　構造方程式は，ε_2，ε_3，ε_4 をそれぞれ平均 0 で互いに独立な誤差項として，以下のようである。

$$X_2 = \alpha_{21}X_1 + \varepsilon_2$$

$$X_3 = \alpha_{32}X_2 + \varepsilon_3$$

$$X_4 = \alpha_{41}X_1 + \alpha_{42}X_2 + \alpha_{43}X_3 + \varepsilon_4$$

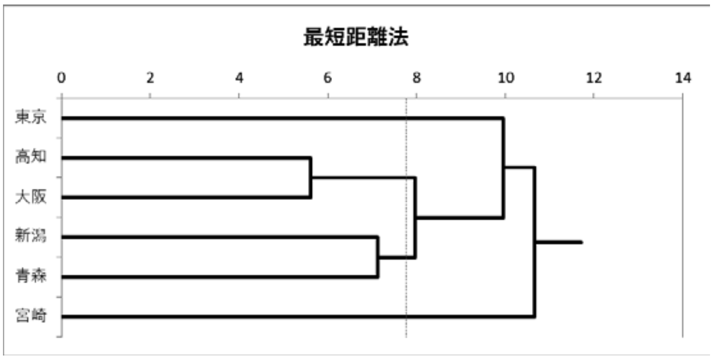
[4]　$\rho_{12} = \alpha_{21}$，$\rho_{13} = \alpha_{32}\alpha_{21}$，$\rho_{23} = \alpha_{32}$，$\rho_{24} = \alpha_{41}\alpha_{21} + \alpha_{42} + \alpha_{43}\alpha_{32}$ であり，$\alpha_{21} = -0.40$，$\alpha_{32} = 0.50$ である。

[5]　総合効果は 0.4 である。

人文４

[1]

(1)　最短距離法を用いたデンドログラムは以下のようである。



(2)　4 つのクラスターに分けられる：

第 1 クラスター：東京　　　第 2 クラスター：高知，大阪

第 3 クラスター：青森，新潟　第 4 クラスター：宮崎

[2]

(1)　鎖効果とは，1 つのクラスターに鎖状にクラスターを形成して行く現象である。最短距離法では，１つでも近いデータがあればクラスターが統合され鎖効果が生じやすい。

(2)　新たに統合されるクラスター内の平方和を最小にする基準でクラスターを形成する。

(3)　ウォード法では，クラスター内のまとまりがよく，クラスター間の距離が適度に保証されるという点で，結果として得られたクラスターの説明がしやすい。

2018.11.25

社会 (18) 略解 (v-F.1)

社会１

[1]　帰無仮説は $H_0 : p_{ij} = p_{i+} p_{+j}$ と表され，H_0 の下での期待度数は $e_{11} \approx 68.2$ である。

[2]　イェーツの補正を行った検定統計量の値は $Y \approx 7.6$ であり，有意水準 0.05 での棄却域は $Y > 3.84$ であるので，帰無仮説は棄却される。

[3]　フィッシャー検定における（片側）P 値は 0.0839 となる。片側 P 値が 0.05 より大きいので両側 P 値も 0.05 よりも大きく，帰無仮説は棄却されない。

[4]　カイ二乗検定はサンプルサイズが大きいときの近似検定であるので，サンプルサイズが小さいときはフィッシャー検定が用いられる。また，フィッシャー検定は周辺度数が与えられた下での条件付き検定であり，カイ二乗検定は無条件での検定との解釈が成り立ち，どちらが望ましいのかについては諸説ある。

社会２

[1]　経済データでは値の大きさに比例して標準偏差が大きくなることがあり，変動係数が妥当な指標となる。変動係数の値の比較から，層 3 での散らばりが最も大きいと結論付けることもできる。

[2]　正規分布に基づく母平均の信頼係数 0.95 の信頼区間に関する考察より公式が得られる。与えられた数値から計算すると $n_1{}^* = 16$ となる。

[3]　$$V[\hat{\mu}] = \sum_{i=1}^{3} \left(\frac{N_i}{N} \right)^2 \frac{N_i - n_i}{N_i - 1} \frac{\sigma_i{}^2}{n_i}$$

[4]　$$n_i{}^{\#} = \frac{\frac{N_i}{N}\sigma_i}{\sum_{i=1}^{3} \frac{N_i}{N}\sigma_i} n$$

[5]　層 1 には 9，層 2 には 20，層 3 には 91 の標本を割り当てる。

社会３

[1]　撹乱項は正規分布よりも裾の重い分布になっていることが読み取れる。

[2]　$V_{t-1}[\varepsilon_t]=E_{t-1}[\varepsilon_t^2]$ であり，$V_{t-1}[\varepsilon_t]=E_{t-1}[h_t]E_{t-1}[\eta_t^2]$ において $E_{t-1}[\eta_t^2]=E[\eta_t^2]=1$ であり，h_t は $t-1$ 期までの情報で決まるので，$E_{t-1}[h_t]=h_t$ となる。

[3]　$V[\varepsilon_t]=\dfrac{\alpha_0}{1-\alpha_1}$

[4]　$E[\varepsilon_t^4]=\dfrac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}$

[5]　ε_t の尖度は $\beta_2=3\left(\dfrac{1-\alpha_1^2}{1-3\alpha_1^2}-1\right)$ となり，$(1-\alpha_1^2)-(1-3\alpha_1^2)=2\alpha_1^2>0$ であるので，ε_t の尖度は必ず正となる。

社会４

[1]　$\sigma > 0$ のとき，mode < median < E[X] である。

[2]　対数尤度関数の微分により最尤推定量は $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \log x_i$ となる。

[3]　μ の事後分布は $N\left(\frac{n\tau^2\overline{y} + \sigma^2\nu}{n\tau^2 + \sigma^2}, \frac{\sigma^2\tau^2}{n\tau^2 + \sigma^2}\right)$ である。

[4]　各推定値は，$E[X]: 2324.5$，median$: 2253.0$，mode$: 2116.5$ と求められる。

[5]　事後期待値は 7.73 である。

[6]　$V\left[\overline{X}\right] = \exp[2\mu + \sigma^2](\exp[\sigma^2] - 1)/n$ および $V\left[\exp[\overline{Y} + \sigma^2/2]\right] \approx \exp[2\mu + \sigma^2]\sigma^2/n$

となる。これらの大小関係はパラメータの値による。

2018.11.25

理工 (18) 略解 (v-F.1)

理工１

[1] $F(x)=\begin{cases}1-e^{-\lambda x} & (x\ge 0)\\ 0 & (x<0)\end{cases}$， $M_X(\theta)=\dfrac{\lambda}{\lambda-\theta}$

[2] 数学的帰納法を用いて示される。また， $M_W(\theta)=\left(\dfrac{\lambda}{\lambda-\theta}\right)^n$ である。

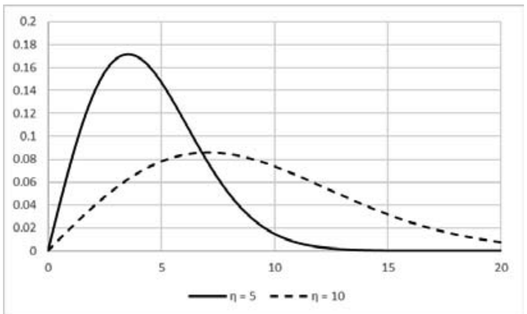
[3] $F(x)=P(X\le x)=P(-\lambda^{-1}\log U\le x)=1-e^{-\lambda x}$

[4] $P(Y=y)=P(y\le X<y+1)=(1-e^{-\lambda})(e^{-\lambda})^y$ となるので， $p=1-e^{-\lambda}$ とすることにより所与の結果が得られる。

[5]
$$M=m \Leftrightarrow X_1+\cdots+X_m<\lambda<X_1+\cdots+X_m+X_{m+1}$$
$$\Leftrightarrow U_1\times\cdots\times U_m>e^{-\lambda}>U_1\times\cdots\times U_mU_{m+1}$$
より示される。

<u>理工２</u>

[1]　$f(x)=\left(\frac{m}{\eta}\right)\left(\frac{x}{\eta}\right)^{m-1}\exp\left[-\left(\frac{x}{\eta}\right)^{m}\right]$。$x_{\text{mode}}=\left(\frac{m-1}{m}\right)^{1/m}\eta$



[2]　$h(x)=\left(\frac{m}{\eta}\right)\left(\frac{x}{\eta}\right)^{m-1}$。$m<1$ では $h(x\,;m,\eta)$ は単調減少，$m=1$ では定数，$m>1$ では単調増加となる。

[3]　システムの寿命 $Y=\min(X_1,\dots,X_k)$ は $W(m,\eta/k^{1/m})$ に従う。

[4]　$P(X<Y)=\dfrac{\eta_2{}^m}{\eta_1{}^m+\eta_2{}^m}$ となる。$m=2$，$\eta_1=10$，$\eta_2=5$ を代入して $P(X<Y)=0.2$ を得る。

[5]　対数尤関数は $l(m,\eta)=n\log(m/\eta)+(m-1)\sum_{i=1}^{n}\log(x_i/\eta)-\sum_{i=1}^{n}(x_i/\eta)^m$ である。また，$\log\log\dfrac{1}{1-F(x)}=-m\log\eta+m\log x$ となるので，$\hat{F}(x)$ を経験分布関数として $(\log x_i,\log\log[1/\{1-\hat{F}(x_i)\}])$, $i=1,\dots,n$ のプロットより各パラメータを算出する（ワイブル確率紙の利用)。

理工３

[1]　$\sigma > 0$ のとき，mode < median < E[X] である。

[2]　対数尤度関数の微分により最尤推定量は $\hat{\mu} = \dfrac{1}{n}\sum_{i=1}^{n} \log x_i$ となる。

[3]　μ の事後分布は $N\left(\dfrac{n\tau^2\bar{y} + \sigma^2\nu}{n\tau^2 + \sigma^2}, \dfrac{\sigma^2\tau^2}{n\tau^2 + \sigma^2}\right)$ である。

[4]　各推定値は，$E[X]: 2324.5$，median : 2253.0，mode : 2116.5 と求められる。

[5]　事後期待値は 7.73 である。

[6]　$V[\bar{X}] = \exp[2\mu + \sigma^2](\exp[\sigma^2] - 1)/n$ および $V\left[\exp[\bar{Y} + \sigma^2/2]\right] \approx \exp[2\mu + \sigma^2]\sigma^2/n$

となる。これらの大小関係はパラメータの値による。

理工４

[1]　S さん用には 2^4 計画の 1/2 実施（2^{4-1} 計画），T さん用には 2^5 計画の 1/4 実施（2^{5-2} 計画）がよい実験計画と言える。

[2]　要因 A の主効果は $(y_1 + y_2 - y_3 - y_4 + y_5 + y_6 - y_7 - y_8)/8$ で求められ，要因 A と要因 B の 2 因子交互作用は $(y_1 + y_2 + y_3 + y_4 - y_5 - y_6 - y_7 - y_8)/8$ で求められる。

[3]　要因 A の主効果は要因 B と要因 E の 2 因子交互作用と交絡する。

[4]　5%有意となるのは要因 A と B である。

[5]　誤差分散を σ^2 とすると，予測分散は，

$$\sigma^2 \boldsymbol{x}'(X'X)^{-1}\boldsymbol{x} = \frac{\sigma^2}{32}\{8 + 5(x_1^2 + x_2^2)^2 - 4(x_1^2 + x_2^2)\}$$

であり，これは $(x_1^2 + x_2^2)$ の関数であることから，計画は回転可能性を有する。

2018.11.25

医薬 (18) 略解 (v-F.1)

医薬１

[1]　尤度関数は $L = \prod_{i=1}^{n} \lambda e^{-\lambda t_i}$ となる。

[2]　尤度関数は $L = \prod_{i=1}^{n} \lambda^{\delta_i} e^{-\lambda t_i}$ となる。

[3]　フィッシャー情報量は $I(\lambda) = \dfrac{\delta}{\lambda^2}$ である。

[4]　漸近分散は $V[\log \hat{\lambda}_1 - \log \hat{\lambda}_2] \approx \dfrac{1}{d_1} + \dfrac{1}{d_2}$ となる。Z 統計量は $Z = \dfrac{\log \hat{\lambda}_1 - \log \hat{\lambda}_2}{\sqrt{\dfrac{1}{d_1} + \dfrac{1}{d_2}}}$ である。

[5]　帰無仮説の下では $d_1 \approx d_2 (= d)$ となるので，検出力から求められる式

$$\frac{|\log \lambda_1 - \log \lambda_2|}{\sqrt{2/d}} = Z_{\alpha/2} + Z_\beta$$

を d について与式が求められる。

医薬２

[1]　分散は $V[\hat{\delta}] = \dfrac{1}{n}(\pi_{12} + \pi_{21} - \pi_{12}{}^2 - \pi_{21}{}^2 + 2\pi_{12}\pi_{21})$ となる。

[2]　δ の近似的な $100(1-\alpha)\%$ 両側信頼区間は，$\hat{V}[\hat{\delta}] = \dfrac{n_{12} + n_{21}}{n^2} - \dfrac{(n_{12} - n_{21})^2}{n^3}$ として

$$\left[\hat{\delta} - Z_{\alpha/2}\sqrt{\hat{V}[\hat{\delta}]}, \hat{\delta} + Z_{\alpha/2}\sqrt{\hat{V}[\hat{\delta}]}\right]$$

となる。

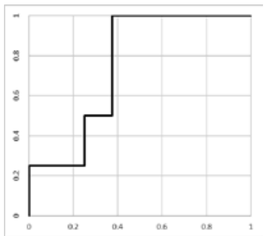
[3]　帰無仮説の下での最尤推定量は $\hat{\pi}_{11} = \dfrac{n_{11}}{n}, \hat{\pi}_{12} = \dfrac{n_{12} + n_{21}}{2n}, \hat{\pi}_{21} = \dfrac{n_{12} + n_{21}}{2n}, \hat{\pi}_{22} = \dfrac{n_{22}}{n}$ となり，対立仮説の下での最尤推定量は $\hat{\pi}_{11} = \dfrac{n_{11}}{n}, \hat{\pi}_{12} = \dfrac{n_{12}}{n}, \hat{\pi}_{21} = \dfrac{n_{21}}{n}, \hat{\pi}_{22} = \dfrac{n_{22}}{n}$ となる。

[4]　尤度比検定統計量は $2\left\{n_{12}\log\left(\dfrac{2n_{12}}{n_{12} + n_{21}}\right) + n_{21}\log\left(\dfrac{2n_{21}}{n_{12} + n_{21}}\right)\right\}$ で与えられ，帰無仮説の下で漸近的に自由度 1 のカイ二乗分布に従う。

医薬３

[1]　真陽性率は $5/8 = 0.625$ であり，偽陽性率は $3/8 = 0.375$ である。

[2]　ROC 曲線は右のようである。



[3]　予測値が 0.35 以上の被験者を疾患 Y に罹患していると判断する。

[4]　求める面積を 2 つの部分に分割し，それぞれの面積を導出する式を組み合わせて示す。

[5]　上問 [4] と同様に面積を分割し，それぞれの面積を導出する式を組み合わせることによって証明する。

医薬４

[1]　対数尤度関数は

$$l(\beta_0, \beta_1, \ldots, \beta_p)$$
$$= \sum_{i=1}^{n} y_i(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}) - \sum_{i=1}^{n} \log(1 + \exp[\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}])$$

である。

[2]　$\log\left(\frac{\theta}{1-\theta}\right) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$ であるので，$x_1 = 0$ あるいは 1 を代入することにより，β_1 が調整オッズ比であることが示される。

[3]　AIC は

$$\mathrm{AIC} = -2\sum_{i=1}^{n} y_i(\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \cdots + \hat{\beta}_p x_{ip})$$
$$+ 2\sum_{i=1}^{n} \log(1 + \exp[\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \cdots + \hat{\beta}_p x_{ip}]) + 2(p+1)$$

となる。

[4]　Kullback-Leibler 情報量は

$$I(g\,; f) = \theta \log\left(\frac{\theta}{\pi}\right) + (1-\theta)\log\left(\frac{1-\theta}{1-\pi}\right)$$

となる。

[5]　リスク因子の同定および AIC の値を踏まえ，ロジスティックモデルは医師単独の予測よりも感度と特異度の観点から優れていることの示唆が得られたことを述べる。

2018.11.21

共通 (18) 略解 (v-F.1)

共通１

[1]　X の期待値 ξ は $\xi=\int_{-\infty}^{\infty}x\cdot\frac{1}{2}\{f_1(x)+f_2(x)\}dx=\frac{\mu_1+\mu_2}{2}$，分散は

$$\tau^2=\int_{-\infty}^{\infty}(x-\xi)^2\cdot\frac{1}{2}\{f_1(x)+f_2(x)\}dx=\sigma^2+\left(\frac{\mu_1-\mu_2}{2}\right)^2$$

となる。

[2]　$\bar{x}=59.65$，$s\approx 12.43$ となる。

[3]　$f(x)$ の 1 次導関数は

$$f'(x)=\frac{1}{2}\cdot\frac{1}{\sqrt{2\pi}\sigma}\left\{-\frac{x-\mu_1}{\sigma^2}\exp\left[-\frac{(x-\mu_1)^2}{2\sigma^2}\right]-\frac{x-\mu_2}{\sigma^2}\exp\left[-\frac{(x-\mu_2)^2}{2\sigma^2}\right]\right\}$$

であり，2 次導関数は

$$f''(x)=-\frac{1}{2}\cdot\frac{1}{\sqrt{2\pi}\sigma}\cdot\frac{1}{\sigma^2}\left\{\left(1-\left(\frac{x-\mu_1}{\sigma}\right)^2\right)\exp\left[-\frac{(x-\mu_1)^2}{2\sigma^2}\right]+\left(1-\left(\frac{x-\mu_2}{\sigma}\right)^2\right)\exp\left[-\frac{(x-\mu_2)^2}{2\sigma^2}\right]\right\}$$

となる。$f'(\xi)=0$ となるので，$x=\xi=\frac{\mu_1+\mu_2}{2}$ は $f(x)$ の極値を与える。

[4]　$f(x)$ が二峰性となるのは，$f(x)$ が $x=\xi$ で下に凸，すなわち $f''(\xi)>0$ のときであるので，条件 $\frac{|\mu_1-\mu_2|}{\sigma}>2$ を得る。