統計解析論試験問題解答例（１部）　　　担当：中川　満　　持ち込み：全て可

１．最小二乗法による残差の和が０になるのが保証されている場合○，保証されていない場合×をつけよ．

（ア）$Y=\beta X+\varepsilon$

（イ）$Y=\alpha+\beta X+\varepsilon$

（ウ）$Y=\alpha X+\beta Z+\gamma W+\varepsilon$

（エ）$Y=\alpha+\beta X+\gamma Z+\delta W+\varepsilon$

ただし，X,Z,W は定数ではない．また，ε は誤差項を表す．

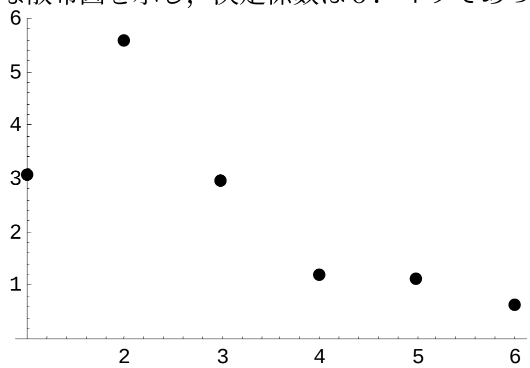
（オ）決定係数が意味をもつのは○の場合か×の場合か答えよ．

（解答例）ア　×，　イ　○，　ウ　×，　エ　○

（解説）

定数項が回帰式に入っていれば，説明変数に１が入っている．この場合，残差と説明変数，１の積和は０になる．残差と１の積は残差そのものであるから，最小二乗残差の合計は０

２．ＸとＹは右図のような散布図を示し，決定係数は０．４９であった．相関係数は？



（解答例）　－０．７

（解説）

相関係数の二乗が決定係数である．問題は，相関係数の符号であるが，それは右下がりの散布図から負と解る．

３．回帰式 $Y=\alpha+\beta X+\gamma Z+\varepsilon$ で，帰無仮説：$\alpha-\beta+\gamma=2$，対立仮説：$\alpha-\beta+\gamma\neq 2$ を検定したい．そのために，回帰式の変形を行う．変形した回帰式を示し，どの説明変数に対する係数についてｔ値を見ればよいのかを示せ．

（解答例）解説を参照のこと

（解説）

＜解答１＞

$$Y=(\alpha-\beta+\gamma-2)+\beta-\gamma+2+\beta X+\gamma Z+\varepsilon$$

$$Y-2=(\alpha-\beta+\gamma)+\beta(X+1)+\gamma(Z-1)+\varepsilon$$

と変形して，定数項のｔ値を見る．

＜解答２＞

$$Y=\alpha+(\alpha-\beta+\gamma-2)X-\alpha X-\gamma X+2X+\gamma Z+\varepsilon$$

$$Y-2X=\alpha(1-X)+(\alpha-\beta+\gamma-2)X+\gamma(Z-X)+\varepsilon$$

と変形して，Ｘの係数のｔ値を見る．

＜解答３＞

$Y=\alpha+\beta X+(\alpha-\beta+\gamma-2)Z-\alpha Z+\beta Z+2Z+\varepsilon$
$Y-2Z=\alpha(1-Z)+\beta(X+Z)+(\alpha-\beta+\gamma-2)Z+\varepsilon$
と変形してＺの係数のｔ値を見る.

その他いろいろな取り方があり得る.

４．説明変数は単位１０億円で計測し, 被説明変数は単位１００億円で計測し, 回帰を行った. その結果，説明変数の係数の推定値は，－０．９，ｔ値は－３．８であった．もし，説明変数と被説明変数の単位を１０億円にそろえたとしたら，係数の推定値はいくつか？ｔ値はどうなるか？また，単位を１００億円にそろえたらどうなるか？

（解答例）

係数推定値	－９．０	－９．０
ｔ値	－３．８	－３．８

（解説）
元の単位での回帰式を$Y=\alpha+\beta X+\varepsilon$とすると，被説明変数の数値は１０倍されるので$10Y=10\alpha+10\beta X+10\varepsilon$である．したがって，係数推定値は１０倍，よって－９．０，ｔ値は変わらず，－３．８．次に，説明変数の単位だけをいじれば，説明変数の数値は１／１０されるので，$Y=\alpha+(10\beta)(X/10)+\varepsilon$となる. 従って, 係数推定値は１０倍になる. つまり, －９．０，ｔ値は変わらない.

５．回帰式$Y=\alpha+\beta X+\varepsilon$を最小二乗法で推定したとき，最小二乗推定量$\hat{\beta}$，$TSS$，$ESS$，$RSS$，$R^2$を計算せよ. ただし$\bar{x}=2$，$\bar{y}=1$，$\sum_{i=1}^{n}y_i^2=30$，$\sum_{i=1}^{n}y_ix_i=10$，$\sum_{i=1}^{n}x_i^2=25$，$n=5$である.

（解答例）$\hat{\beta}=0, TSS=25, ESS=0, RSS=25, R^2=0$

（解説）
新作問題集と違って定数項が入っていることに注意.

$$\hat{\beta}=\frac{\sum_{i=1}^{n}x_iy_i-n\bar{x}\,\bar{y}}{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}=\frac{10-5\times2\times1}{25-5\times2^2}=0$$

$$TSS=\sum_{i=1}^{n}y_i^2-n\bar{y}^2=30-5\times1^2=25$$

$$ESS=\hat{\beta}^2\left(\sum_{i=1}^{n}x_i^2-n\bar{x}^2\right)=0$$

$$RSS=TSS-ESS=25$$

$$R^2=ESS/TSS=0$$

６．回帰式$Y=\alpha+\beta X+\gamma Z+\delta W+\varepsilon$において$\beta$に関するｔ値が－１．５であった，このとき, このＹとＸのＺ, Ｗの影響を取り除いた偏相関係数$r_{X,Y|Z,W}$を求めよ. ｎ＝１４とする.

（解答例）$r_{X,Y|Z,W}=-0.428571$

（解説）

$$r^2_{X,Y|Z,W} = \frac{t^2}{n-K+t^2} = \frac{(-1.5)^2}{14-4+(-1.5)^2} = 0.183673$$

$$r_{X,Y|Z,W} = \operatorname{sgn} t\sqrt{0.183673} = -0.428571$$

７．被説明変数Ｙに対して考え得る説明変数としては，Ｘ，Ｚ，Ｖ，Ｗの４つと定数が考えられる．説明変数を定数，Ｘ，Ｚ，Ｖ，Ｗの順に入れていくとして，以下のようなＲＳＳが得られた．自由度修正付き決定係数によってモデル選択をして，回帰式に含めるべき説明変数を決めよ．

ｎ＝１５，ＴＳＳ＝６０

説明変数	RSS
定数	40
定数，X	30
定数，X，Z	20
定数，X，Z，V	19
定数，X，Z，V，W	18

（解答例）$Y = \alpha + \beta X + \gamma Z + \varepsilon$

（解説）

$\bar{R}^2 = 1 - \dfrac{RSS/(n-K)}{TSS/(n-1)}$ によって自由度修正付き決定係数を計算すると，

説明変数	RSS	
定数	40	0.333
定数，X	30	0.462
定数，X，Z	20	0.611
定数，X，Z，V	19	0.597
定数，X，Z，V，W	18	0.580

となるので，$Y = \alpha + \beta X + \gamma Z + \varepsilon$ が選択される．

８．被説明変数Ｙに対して考え得る説明変数としては，Ｘ，Ｚ，Ｖ，Ｗの４つと定数が考えられる．説明変数を定数，Ｘ，Ｚ，Ｖ，Ｗの順に入れていくとして，新たに追加した説明変数に対するｔ値はそれぞれ，１．５，－１．４，１．９，－０．８であった，この場合，自由度修正付き決定係数によってモデル選択をして，回帰式に含めるべき説明変数を決めよ．

（解答例）$Y = \alpha + \beta X + \gamma Z + \delta V + \varepsilon$

（解説）

自由度修正付き決定係数が増えるのは新しく回帰に入れた説明変数のｔ値の絶対値が１より大きい場合である．したがって，Ｘ，Ｚ，Ｖまではｔ値の絶対値が１より大きいので，自由度修正付き決定係数が増加するが，Ｗをいれた時点で，自由度修正付き決定係数は減少する．

９．以下のような回帰を行ったらｔ値が異常小さくなってしまった．原因を述べよ．

$Y = \alpha + \beta_1 D_1 + \beta_2 D_2 + \gamma X + \varepsilon$．ただし，

$$D_1 = \begin{cases} 1 & \text{その月が１～６月の場合} \\ 0 & \text{それ以外} \end{cases}$$

$$D_2 = \begin{cases} 1 & \text{その月ががが7～12月の場合} \\ 0 & \text{それ以外} \end{cases}$$

とする．

（解答例）多重共線性の発生

（解説）$D_1 + D_2 = 1 =$ 定数項の説明変数だから．

１０．以下は米国の１９１０年から１９８８年までのデータを元に実際に回帰分析を最小二乗法で行った時の，EXCEL の分析ツールの出力結果である．これをみて以下の問いに答えよ．

回帰モデル：

$$\text{実質株価収益率} = \alpha\, logrealgnp + \beta\, realinterestrate + \gamma\, unemployment\ rate + \delta\, gnpgrowth + \phi\, price + \eta + \text{誤差項}$$

ただし，logrealgdp：実質 GNP（log 変換後），realinterestrate:実質利子率，unemployment rate:失業率，gnpgrowth:実質ＧＮＰ成長率，price:物価水準

（EXCEL の出力［() 内は教科書での表現]）

概要

回帰統計

重相関 R（$\sqrt{R^2}$）	XXXXXXXX
重決定 R2（R^2）	（ア）
補正 R2（自由度修正 R^2）	（イ）
標準誤差（$S = \sqrt{S^2}$）	（ウ）
観測数（ｎ）	79

分散分析表

	自由度	変動	分散	観測された分散比	有意 F
回帰	５	6724.472981 （回帰変動）	XXXXXX （回帰変動/自由度）	6.075323423 （全係数 F 検定）	9.32772E-05 （そのｐ値）
残差	（エ）	16160.013 （残差変動）	（オ） （$\frac{1}{n-K}\sum_{i=1}^{n}\hat{u}_i^2$）		
合計	（カ）	22884.48598 （総変動）			

	係数	標準誤差	t	P-値	下限 95%	上限 95%
	（係数推定値）	（$S/\sqrt{\sum_{i=1}^{n}\left(x_i^*\right)^2}$）	（t 統計量）	（P 値）	（95%信頼区間）	
切片	−0.117317323	17.21847452	（キ）	0.995	（ク）	（ケ）
logrealgnp	−4.819619872	7.783365217	（コ）	0.538	−20.3	10.7
realinstrest	1.360349722	0.354749269	（サ）	0.000	0.653	2.07
unemployment	−0.724281459	0.362795972	（シ）	0.0496	−1.45	−0.00123
gnpgrowth	1.344659003	0.312243469	（ス）	0.000	0.722	1.97
price	6.483028379	7.782368899	（セ）	0.408	-9.03	22.0

（１）（ア）から（セ）までの欄に数値を与えよ.

（ア）0.29　（イ）0.25　（ウ）14.9　（エ）７３　（オ）221.4

（カ）７８　（キ）−0.01　（ク）−34.4　（ケ）34.2　（コ）-0.62

（サ）**3.8**　（シ）−2.0　（ス）4.3　（セ）0.83

（２）（ソ），（タ），（チ）について，検定統計量値，境界値，検定結果を示せ.

（ソ）対立仮説 $\beta > 0$，帰無仮説 $\beta = 0$ を有意水準５％で検定すると帰無仮説は棄却されるか？

検定統計量値：3.8　境界値：1.65　検定結果：棄却される

（タ）対立仮説 $\gamma < 0$，帰無仮説 $\gamma = 0$ を有意水準５％で検定すると帰無仮説は棄却されるか？

検定統計量値：−2.0　境界値：−1.65　検定結果：棄却される

（チ）対立仮説帰 $\phi \neq 0$，無仮説 $\phi = 0$ を有意水準１％で検定すると帰無仮説は棄却されるか？

検定統計量値：0.83　境界値：±2.58　検定結果：棄却されない