統計解析論模擬試験問題

（実際の試験はこの１／４位の量の予定です．それとあくまでも類似の問題がでるだけですから，試験対策を怠りなくお願いします.）

１．以下の文の（）内を下の記号式群から最も適当なものを選んで埋めよ．重複がありうるので注意すること.

回帰式　消費＝α＋β所得＋誤差項において，説明変数は（ア）と定数で被説明変数は（イ）である．これを（ウ）で推定するというのは，{(エ）のデーター（α＋β×所得のデータ)}の二乗を全データについて合計したものを（オ）にする（カ），（キ）の値を推定値とするものである．この回帰式の意味は，消費＝α＋β所得という（ク）の関係が近似的に成立するというものである.

一般的には，（ケ）をY，説明変数をXとかく．（コ）に関するデータについては，i 番目のデータという意味でy_i，（サ）についてはx_iと書く．このとき回帰式を$y_i = \alpha + \beta x_i + \varepsilon_i$と書く．$\varepsilon_i$は（シ）である．このとき，最小二乗法で推定すると，βの推定量は$\hat{\beta} =$（ス），αの推定量は$\hat{\alpha} =$（セ）である．ここで，$\bar{x}$は（ソ），つまり，X のデータの（タ），$\bar{y}$は（チ），つまり，Y のデータの（ツ）である.

＜記号式群＞

一次式, 誤差項, 最小, 最小二乗法, 消費, 所得, 説明変数, 被説明変数, 標本平均, $\frac{1}{n}\sum_{i=1}^{n} x_i$,

$\frac{1}{n}\sum_{i=1}^{n} y_i$ ，α，β，$\frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$ ，$\bar{y} - \hat{\beta}\bar{x}$

２．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 （係数推定値）	標準誤差 （$S/\sqrt{\sum_{i=1}^{n}\left(x_i^*\right)^2}$）	t （t 統計量）	P-値 （P 値）	下限 95% （95%信頼区間）	上限 95%
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.

（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？

（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？

（チ）対立仮説帰 $\phi \neq 0$，無仮説 $\phi = 0$ を有意水準５％で検定すると帰無仮説は棄却されるか？以下の文章の空欄を補え．ただし，ト，ニについては，{棄却される，棄却されない} のいずれから選び，それ以外は数値を補え．

回帰モデル

$$\text{実質株価収益率} = \beta\ realinterestrate + \gamma\ unemployment\ rate + \delta\ gnpgrowth + \eta + \text{誤差}$$

における残差変動は，16385.92 であった．このとき，帰無仮説 $\alpha = \phi = 0$ を有意水準５％で検定したとき，そのＦ統計量は（ツ）であり，攪乱項が正規分布で独立同分布だとしたときの統計量の帰無仮説の下でのＦ分布の自由度は２，(エ) であり，その５％の境界値（臨界値）は（テ）である．従って，帰無仮説は（ト）．なお，この回帰式の係数の少なくとも１つは０でないという対立仮説を（ナ）で検定すると帰無仮説は（ニ）．また，実質 GNP（log 変換後）の偏相関係数は（ヌ）である．

＜数表＞

ｔ分布表

自由度	p 0.05	p 0.025
70	1.666915	1.994435
71	1.666599	1.993944
72	1.666294	1.993462
73	1.665996	1.992998
74	1.665708	1.992544
75	1.665426	1.992103
76	1.665151	1.991675
77	1.664885	1.991257
78	1.664625	1.990848
79	1.664371	1.990452

Ｆ分布表

分母自由度	分子自由度 1	2	3
70	3.977789	3.127681	2.73554
71	3.975813	3.125763	2.733643
72	3.973895	3.123901	2.731809
73	3.972048	3.122103	2.730019
74	3.970229	3.120348	2.728278
75	3.968466	3.118643	2.726594
76	3.966761	3.11698	2.724946
77	3.965084	3.11536	2.72334
78	3.963464	3.113797	2.721784
79	3.961901	3.112262	2.720263

３．以下の文章の空欄を下の語式群で補い，その記号を答えよ．なお，同一の語式が複数の空欄を埋めうるとする．なお語式群の記号はＤ，Ｊ，Ｖの文字を含まないようになっているので注意すること．

「回帰モデルを $y_i = \alpha + \beta x_i + \varepsilon_i \quad (i = 1, \cdots, n)$ とする．」という文章は，二つの意味を持っている．一つは，（ア）y_i のデータが（イ）x_i の一次式 $\alpha + \beta x_i$ と誤差項 ε_i の和で生成されているというもので，もう一つの意味は，$y_i = \alpha + \beta x_i$ +誤差というモデルをたてて．何らかの方法で α, β の値を推定しようというものである．前者は，データの成り立ち（DGP）を説明したもの

で，後者は，一次式のモデルに当てはめるというモデルを説明したものである．一般には，上記の一言でモデルと DGP の間に矛盾がないケースを考えるということを表す．このモデルの推定に（ウ）を使用するとは，$\sum_{i=1}^{n}\left(y_i-a-bx_i\right)^2$ を最小化する a,b をそれぞれ α,β の推定値とすることを意味している．このとき α,β の推定量を $\hat{\alpha},\hat{\beta}$ とすると，$\hat{\beta}=$（エ），$\hat{\alpha}=$（オ）となる．ただし，$\bar{x}$ はｘの（カ），$\bar{y}$ はｙの（キ）である．このとき，回帰残差 $\hat{u}_i=$（ク）を求めると，$\sum_{i=1}^{n}\hat{u}_i=$（ケ）である．また，$\sum_{i=1}^{n}\hat{u}_i x_i=$（コ）である．さらに，このような定数項を含む回帰モデルでは，全変動<TSS>（(サ)）=回帰変動<ESS>（(シ)）+残差変動<RSS>（(ス)）が成り立つので，回帰の当てはまりの指標として被説明変数の変動のうちどれだけの割合が説明変数の変動で説明できるかを示す（セ）$R^2=$（ソ）を考えることができる．これが大きいほど当てはまりは（タ）．また，回帰式に定数項がある場合は，（チ）$\le R^2\le$（ツ）．実は，ｘとｙの重相関係数（テ）の二乗は（ト）と一致する．したがって，重相関係数による分析は定数項とｘを説明変数とした線形回帰分析とまったく同じことをやっているのである．

（語式群）

（A）標本平均，(B)標本分散，(C)決定係数，(E)相関係数，(F)被説明変数，(G)説明変数，(H)最小二乗法，(I) よい，(K) 悪い，(L) $y_i-\hat{\beta}x_i$，(M) $y_i-\hat{\alpha}-\hat{\beta}x_i$，(N) $\bar{y}-\hat{\beta}\bar{x}$，(O) $\bar{y}+\hat{\beta}\bar{x}$，(P) $y_i+\beta x_i$，(Q) $\sum_{i=1}^{n}x_i^2$，(R) $\sum_{i=1}^{n}y_i^2$，(S) $\sum_{i=1}^{n}\hat{u}_i^2$，(T) $\sum_{i=1}^{n}x_i y_i$，(U) $\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$，(W) $\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2$，(X) $\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)$，(Y) $\sum_{i=1}^{n}\hat{\alpha}^2\left(x_i-\bar{x}\right)^2$，(Z) $\sum_{i=1}^{n}\hat{\beta}^2\left(x_i-\bar{x}\right)^2$，(AA) $\sum_{i=1}^{n}\hat{\alpha}^2\left(y_i-\bar{y}\right)^2$，(BB) $\sum_{i=1}^{n}\hat{\beta}^2\left(y_i-\bar{y}\right)^2$，(CC) $\frac{ESS}{RSS}$，(EE) $\frac{ESS}{TSS}$，(FF) $\frac{RSS}{ESS}$，(GG) $\frac{RSS}{TSS}$，(HH) $\frac{TSS}{ESS}$，(II) $\frac{TSS}{RSS}$，(KK) $\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\sqrt{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}}$，(LL) $\frac{\sum_{i=1}^{n}x_i y_i}{\sum_{i=1}^{n}x_i^2}$，(MM) $\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}$，(NN) ０である，(OO) ０ではない，(PP)－２，(QQ)－１，(RR)０，(SS)１，(TT)２

４．　以下の文章の空欄を下の語式群で補い，その記号を答えよ．なお，同一の語式が複数の空欄を埋めうるとする．なお語式群の記号はＤ，Ｊ，Ｖの文字を含まないようになっている

ので注意すること．(各２点，計 30 点)

回帰モデルを $y_i = \alpha + \beta x_i + \varepsilon_i \ \ (i = 1, \cdots, n)$ とする．ただし，$E(\varepsilon_i) = 0$，$V(\varepsilon_i) = \sigma^2$，$\varepsilon_i$ は i に関して独立とする．β の推定量を $\hat{\beta}$ とすると，$V\left[\hat{\beta}\right] =$（オ）．なお，$\varepsilon_i$ が平均０，分散 σ^2 の正規分布に従えば，$\hat{\beta}$ は平均（セ），分散（オ）の（ソ）((K)正規分布）である．

（語式群）

(A)ある，(B)ない，(C)一致性，(E)不偏性，(F)漸近正規性，(G)最良性，(H)ｔ分布，(I)Ｆ分布，

(K)正規分布，(L) α，(M) β，(N) $\alpha + \beta$，(O) $\frac{1}{24}$，(P) $\frac{1}{2}n(n+1)$，(Q) $\frac{1}{12}n(n+1)(n-1)$，

(R) $\frac{1}{6}n(n+1)(2n+1)$，(S) $V(\hat{\beta}) \to 0$，(T) $V(\hat{\beta}) \to C < \infty$，(U) $V(\hat{\beta}) \to \infty$，

(W) $\frac{1}{\bar{x}}\sum_{i=1}^{n}(x_i - \bar{x})^2 \longrightarrow 0$，(X) $\frac{1}{\bar{x}}\sum_{i=1}^{n}(x_i - \bar{x})^2 \longrightarrow K < \infty$，(Y) $\frac{1}{\bar{x}}\sum_{i=1}^{n}(x_i - \bar{x})^2 \longrightarrow \infty$，

(Z) $\sum_{i=1}^{n}(x_i - \bar{x})^2 \longrightarrow 0$，(AA) $\sum_{i=1}^{n}(x_i - \bar{x})^2 \longrightarrow K < \infty$，(BB) $\sum_{i=1}^{n}(x_i - \bar{x})^2 \longrightarrow \infty$，

(CC) $\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$，(EE) $\frac{n\bar{x}\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$，(FF) $\frac{2\bar{x}\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$，(GG) $\frac{\sigma^2\sum_{i=1}^{n}x_i^2}{\left\{\sum_{i=1}^{n}(x_i - \bar{x})^2\right\}^2}$，

(HH) $\frac{\sum_{i=1}^{n}x_i\varepsilon_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$，(II) $\frac{\sum_{i=1}^{n}(x_i - \bar{x})\varepsilon_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$，(KK) $\frac{\sum_{i=1}^{n}x_i E[\varepsilon_i]}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$，(LL) $\frac{\sum_{i=1}^{n}(x_i - \bar{x})E[\varepsilon_i]}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$，

(MM) $\sum_{i=1}^{n}\left(\frac{x_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^2 V[\varepsilon_i]$, (NN) $\sum_{i=1}^{n}\left(\frac{(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^2 V[\varepsilon_i]$, (OO) $\sum_{i=1}^{n}\frac{x_i}{\sum_{i=1}^{n}(x_i - \bar{x})^2}V[\varepsilon_i]$,

(PP) $\sum_{i=1}^{n}\frac{(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}V[\varepsilon_i]$,

５．以下の文章の空欄を下の語式群で補い，その記号を答えよ．なお，同一の語式が複数の空欄を埋めうるとする．なお語式群の記号はＤ，Ｊ，Ｖの文字を含まないようになっているので注意すること．

回帰モデルを $y_i = \alpha x_i + \beta z_i + \varepsilon_i \quad (i = 1, \cdots, n)$ とする．ただし，$E(\varepsilon_i) = 0$，$V(\varepsilon_i) = \sigma^2$，$\varepsilon_i$ はｉに関して独立とする．$\hat{\alpha}$ は以下のように求められる．まず，x_i を z_i に回帰する．このときの係数の最小二乗推定量は（ア）であるので，この回帰残差 x_i^* ＝（イ）を求め，y_i を x_i^* に回帰する．その結果，$\hat{\alpha}$ ＝（ウ）＝（エ）となる．また，$\hat{\alpha}$ の分散は，（オ）である．新たな変数を回帰に加えると（カ）は減少するので，（キ）は必ず増加する．もし，回帰式に定数項が含まれていれば，（キ）は意味を持つが，上記から変数を回帰式に加えることの妥当性を示していない．そのため，自由度修正付き（キ）が当てはまりの指標として与えてられている．それは上記の回帰式の場合，式（ク）で与えられる．また，σ^2 の（ケ）推定量は，自由度を考えて S^2 ＝（コ）である．

（語式群）

(A) 不偏，(B)一致，(C)最良，(E)決定係数，(F)相関係数，(G)ESS,(H)TSS,(I)RSS，

(K) $\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i^2$，(L) $\frac{1}{n-1}\sum_{i=1}^{n}\hat{u}_i^2$，(M) $\frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2$，(N) $\frac{ESS/n}{RSS/n}$，(O) $\frac{ESS}{RSS/(n-2)}$，

(P) $1-\frac{RSS/n}{TSS/n}$，(Q) $1-\frac{RSS/(n-2)}{TSS/(n-1)}$，(R) $\frac{\sum_{i=1}^{n}x_i z_i}{\sum_{i=1}^{n}z_i^2}$，(S) $\frac{\sum_{i=1}^{n}y_i x_i^*}{\sum_{i=1}^{n}(x_i^*)^2}$，(T) $x_i - \frac{\sum_{i=1}^{n}x_i z_i}{\sum_{i=1}^{n}z_i^2}z_i$，

(U) $\frac{\sum_{i=1}^{n}z_i^2\sum_{i=1}^{n}x_i y_i - \sum_{i=1}^{n}x_i z_i\sum_{i=1}^{n}y_i z_i}{\sum_{i=1}^{n}x_i^2\sum_{i=1}^{n}z_i^2 - \left(\sum_{i=1}^{n}x_i z_i\right)^2}$，(W) $\frac{\sigma^2}{\sum_{i=1}^{n}\left(x_i - \frac{\sum_{i=1}^{n}x_i z_i}{\sum_{i=1}^{n}z_i^2}z_i\right)^2}$

６．() 内を最も適当な語句，数式で埋めよ．

残差と説明変数の（ア）の合計は０である．説明変数として（イ）がある場合は残差の合計が０になる．しかし，（イ）がない場合は，残差の合計が０になるとはかぎらない．このことから，説明変数に（イ）が含まれていない場合，（ウ），自由度修正付き（ウ）は意味

を持たない．

残差分散は（エ）の分散の推定量である．残差標準偏差はその（オ）である．残差を残差標準偏差で割ったものを（カ）と呼ぶ．（カ）が２を超える確率は少ないと考えられるので，これを超える残差の個数が全標本数の１０％程度を越えている場合，回帰分析が成功しているとはいえないと見なせる．また，特定のカテゴリーのデータ（時系列データの場合特定の期間，クロスセクションデータの場合特定の地域など）に（カ）が２を超えるのが集中している場合，このカテゴリーについては回帰係数が異なる，回帰は当てはまっていないと考えられる．

(語群)

決定係数，誤差項，積，定数項，標準化残差，平方根，

７．以下の Excel の分析ツール（回帰分析）の結果を見て以下の問に答えよ．

	係数	標準誤差	t	P-値
切片	－269858.6398	10180637861	－2.6507E-05	0.99997888
GDE(=GDP)	0.849087251	10180.63786	8.34022E-05	0.999933546
変数２	0.275	10180.63786	2.70121E-05	0.999978477

この結果から想定されるのはなにか？

８．（ア），（イ）に式を補え．

回帰式 $y_i = \alpha + \beta x_i + \gamma z_i + \varepsilon_i$ において帰無仮説 $\beta - \gamma = 1$ を F 検定を使わずに検定したい．そのために，まず，x_i の係数が $\beta - \gamma - 1$ となるように回帰式の和差操作で変形すると，$y_i - (ア) = (\beta - \gamma - 1)x_i + (イ) + \varepsilon_i$ となる．この回帰では、（ウ）を x_i と（エ）と定数に回帰して，（オ）の係数の t 統計量を求めて，対立仮説が $\beta - \gamma \neq 1$ なら（カ）側検定，対立仮説が $\beta - \gamma < 1$ か対立仮説が $\beta - \gamma > 1$ の場合は（キ）側検定で検定を行う．

９．クロスセクション分析で注意しなければならないのは，サイズ効果による見せかけの相関である．これは，地域なり企業の大きさ（人口，面積，資本金，売り上げなど）の要因によって本来は関係がないものが回帰では（ア）があるように見えてしまうことである．時系列データで問題になるのは，本来無関係の増加関数どうし，または，減少関数でも（イ）があるように見えて

しまうことである．この場合は，それぞれはタイムトレンドを持っているにすぎず，本当に両者の間に関係があるなら，タイムトレンドからの乖離に相関があるはずである．この場合は，説明変数にタイムトレンドを加えて回帰し，（ウ）やｔ値を見るべきである．

（語群）
相関，偏相関係数

１０．＜これはそのまま試験に出します＞
「〈子〉のつく名前の女の子は頭がいい」というタイトルの本の中で，「女子高校の偏差値の高さとその高校内の子の付く名前の女性の比率が正の相関を持っている」という事実があることが触れられています．この事実はタイトルの主張する命題の証明となるでしょうか？回帰分析と関連させながら述べてください．