問題2
推定に関する以下の問いに答えよ

(1)「推定値」と「推定量」の違いを簡潔に説明せよ
母数を推定するために得られた統計量が「推定量」で、その推定量の実現値が「推定値」。

(2)「最小二乗推定」と「最尤推定法」の相違を簡潔に説明せよ
ある標本群が与えられているとする。
その標本群は本来何らかのモデル式に従うものであるとして、実現値とモデル式の値の差は誤差であると考える。その誤差が正規分布に従うとし、誤差が最小となるようにモデル式の母数を決定するのが「最小二乗推定」である。
その標本群の実現値が現れたのはそれが最も起こりやすいことであったからと考える推定法で、各標本が実現する同時確率が最大になるように母数を推定するのが「最尤推定法」である。
(要するに全然違うので、どこがどう違うとかいう説明はできません)

(3)
標本 $X=(x_1,x_2,\cdots,x_n)$ が正規分布$N(\mu,\sigma^2)$に従うとき、
その母数、平均μ、分散σ^2を最尤推定法で推定する。

$X=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(X-\mu)}{2\sigma^2}\right)$ より

尤度 ： $L(\mu,\sigma,X)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x_i-\mu)^2}{2\sigma^2}\right\}$

対数尤度 ： $l(\mu,\sigma,X)=n\cdot\ln\frac{1}{\sqrt{2\pi}\sigma}-\sum_{i=1}^{n}\left\{\frac{(x_i-\mu)^2}{2\sigma^2}\right\}$

これが推定量$\hat{\mu},\hat{\sigma}^2$で最大となるので

$$\frac{\partial}{\partial\mu}l(\mu,\sigma,X)|_{\mu=\hat{\mu}}=\frac{(x_i-\hat{\mu})}{\sigma^2}=0$$

$$\frac{\partial}{\partial\sigma}l(\mu,\sigma,X)|_{\sigma=\hat{\sigma}}=\sum_{i=1}^{n}\left\{\frac{(x_i-\mu)^2}{2\sigma^4}\right\}-\frac{n}{2\sigma^2}=0$$

これを解くと

$$\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2$$

となり、それぞれ標本平均、標本分散となっている。

平成 17 年度　数理分析手法 I　試験問題

（注 1）試験中に講義時配付資料を閲覧することを禁じる．

（注 2）関数電卓およびパソコンの使用を認める．パソコン使用の場合はエクセルや関数電卓機能などの使用のみを認め，インターネットの接続は当然のことながら禁じる．諸君の倫理観のある対応を期待する．

問題 1 (25 点)

統計学に関する以下の用語をそれぞれ 2 行程度で説明せよ．

(1)ベイズの定理

(2)中心極限定理

(3)標本分布

(4)推定量

(5)有意水準

問題 2 (20 点)

推定に関する以下の問題に答えよ．

(1) 最小二乗推定と最尤推定の相違を簡潔に説明せよ．

(2) 標本 $\mathbf{X} = (x_1, x_2 \cdots, x_n)$ が正規分布 $N(\mu, \sigma^2)$ からの観測値であるとする．母数を最尤推定法で求めると，それぞれ標本平均と標本分散になることを示せ．

問題 3 (25 点)

ある交差点の月間事故件数を 5 年間(60 ヶ月)調べたところ表の通りであった．月間事故件数がポアソン分布に従っているという仮説を検定したい．以下の問題に答えよ．

月間事故件数	0	1	2	3	4	5
観測頻度	5	12	24	12	6	1

$p = \frac{\lambda^r}{r!}\exp(-\lambda)$

(1) 月間事故件数の理論確率を表す式を記せ．

(2) 行うべき仮説検定の種類と，その時の帰無仮説 H_0 と対立仮説 H_1 を記せ．

(3) 仮説検定を実施せよ．有意水準は 0.05 とする．検定のための分布表は別途配付する．

(2)
主成分分析においては、新たに作る総合的指標によって、各標本間((1)でいうテストを受けた生徒間)の差がはっきりと出る方が良いので、L の分散は大きい方がよい。
分散共分散行列 S とすると、L の分散は

$$E(L) = l^t S l$$

また l は 条件: $l^t l = 1$ より、標本の分散共分散行列 S の固有ベクトルとなる。
このとき固有値λとすると

$$Sl = \lambda l, \qquad E(L) = \lambda$$

となるので、固有値が大きいほどよい主成分であるということになる。
つまり、ある主成分によるばらつきがすべての主成分(n個ある:nは変量の数)によるばらつきのうちどの程度の割合を占めているのかという指標(寄与率)が

$$k_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}$$

であらわすことができる。

(＊)
標準化したデータを用いた場合、分散共分散行列の代わりに相関行列をもちいることができ、

このとき $\sum_{j=1}^{n} \lambda_j = n$ である。

つまり固有値が1以上という条件は、寄与率が $k_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} = \frac{1}{n}$ 以上であるということであるので、全n個の主成分のうち平均以上の寄与率をもつ主成分を選ぶということである。

(テストの解答としてはたぶん(＊)以下で大丈夫。てか別に教科書とかなくて、レジュメ見て式を見ての俺なりの解釈なので正解かどうかわかりません。ちょっとググったくらいでは「何故1以上なのか」までは教えてくれませんでした。。。笑
&よく考えたらすべてにおいてそうなので、このどの解答もあってるかどうかわかりません笑)

(4)
不偏性(unbiased)
一致性(consistent)
有効性(efficient)
十分性(sufficient)　　　のうち3つ

問3
多変量解析について以下の問いについて答えよ

(1)重回帰分析と主成分分析の違いを端的に説明せよ
ある標本iについて、複数の変量がある。
「重回帰分析」はそのうちある一つの変量(目的変量)を他の変量(説明変量)を使って表すことで分析する方法で、「主成分分析」は変量を圧縮して新たな総合的指標をつくり分析する方法。例えば、英語、数学、国語、理科、社会の5教科のテストの得点(変量)があるとき、英語の得点を他の4教科の得点の関数として表すのが「重回帰分析」で、「理系科目の力」などの新しい総合的指標で標本を分析するのが「主成分分析」。

(具体的には
「重回帰分析」は
目的変量 ： Y　説明変量 ： $X_1 \sim X_n$　　残差 ： e　　としたとき

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + e$$

で表わされるとして、残差の二乗和が最小となるようにβを定める。

「主成分分析は」
特性値(総合的指標) ： L　　説明変量 ： $X_1 \sim X_n$

$$L = l_1 X_1 + l_2 X_2 + \cdots + l_n X_n$$ を

$l_1{}^2 + l_2{}^2 + \cdots + l_n{}^2 = 1$ の条件下で L の分散が大きくなる方から順番に、かつ l が互いに相関をもたないように l を定める。)

平成 18 年度　数理分析手法Ⅰ　試験問題

（注 1）試験中に講義時配付資料を閲覧することを禁じる．
（注 2）解答用紙は２枚配付する．
（注 3）この試験は 70 点満点である．

問題 1(20 点)
以下の統計用語を簡潔に説明せよ．式は用いても構わない．
(1)有意水準
(2)中心極限定理 → p14
(3)ベイズの定理 → p12
(4)標本分布 → p15
(5)自由度 → p16

問題 2 (30 点)
推定に関する以下の問題に答えよ．
(1)「推定値」と「推定量」の違いを簡潔に説明せよ． → p18
(2)「最小二乗推定」と「最尤推定」の相違を簡潔に説明せよ．
(3) 標本 $\mathbf{X}=(x_1,x_2\cdots,x_n)$ が正規分布 $N(\mu,\sigma^2)$ からの観測値であるとする．その母数を最尤推定法で求めると，それぞれ標本平均と標本分散になることを示せ．
(4)信頼性の高いデータ分析のために推定量が具備すべき特性を３つ列挙せよ．

問題 3 (20 点)
多変量解析に関する以下の問題に答えよ．
(1)重回帰分析と主成分分析の違いを端的に説明せよ．式や図を用いても構わない．
(2)標準化データを用いた主成分分析において，主成分数を決定する時の考え方として，固有値が 1 以上という条件があるが，その意味するところは何か，説明せよ．