指数分布からの truncated samples に基いた 寿命試験に就いて

(On the Life-Testing Based on the Truncated Samples from the Exponential Distridutions)

田村亮二

§ 1. 序

k個の母集団—例えば，k個の異る方法でつくつた電球の寿命，k個の異る品種の穀物の生産高，等々——を特徴付けているある未知常数について一様性を検定するという問題は普通の実験に於いて，屢々経験する所である。之に対ては，Fisher の輝かしい分散分析法があり又 Bartlett のL-検定法がある。然しいづれも正規母集団を仮定し，指数分布がより合理的と思われる寿命試験については適用されにくい点もある。更に寿命試験では，各仕切からの任意標本をそのまゝ，用いないで中途で打切つた data を使用する場合が多い。此の小論に於いて次の如き寿命試験に現われる一様性検定の問題を取扱ふ。即ち密度函数（今後 *p. d. f.* と略記す）が $f(x_i\, A_i, \theta_i)$ $i=1, 2, \cdots\cdots k$ であるk個の population からの各々大きさ $n_1, n_2 \cdots\cdots, n_k$ の ordered samples を $x_{11} \leqq x_{12} \leqq \cdots\cdots \leqq x_{1n_1}$; $x_{21} \leqq x_{22} \leqq \cdots\cdots \leqq x_{2n_2}$; $\cdots\cdots$; $x_{k1} \leqq x_{k2} \leqq \cdots\cdots \leqq x_{kn_k}$ とし之を夫々初めの $r_1, r_2, \cdots\cdots, r_k$ 個で打切つた truncated samples を基にして A_i, θ_i の一様性に関する統計的仮説を検定する手続を考察する。

$$(1)\quad f(x_i\, A_i,\ \theta_i) = \begin{cases} \dfrac{1}{\theta_i}\, exp\left[-\dfrac{1}{\theta_i}(x-A_i)\right] & x \geqq A \\ O & \text{その他} \end{cases}$$

但し，$A,\ \theta > O$ とす。

此時 x は各 item の寿命の長さを表わす変数であり，A 及び θ は夫々 minimum life 及び分散を表わす常数である。指数分布に関する若干の性質は〔1〕で又 two population の場合は〔2〕に於いて取扱れている。こゝでは〔2〕同様 likelihood ratio method を採用し，一般に広く知られているカイ自乗及び Snedecor の F 分布による検定に帰着させることに成功した。先ず §2 で統計的仮説，test statistics 及びその分布をのべ §3 で test statistics の分布を誘導するための予備定理を証明し，§4 で検定方式の解析を行う。

§ 2. 統計的假説及び検定方式

先づ問題になる仮説及びそれに対する検定の手続をまとめておく。仮説として考えられるものは，

(A) θ に関する仮説

(a) $H_1: \theta_1 = \theta_2 = \cdots\cdots = \theta_k$（但 A_i $i=1, 2, \cdots\cdots, k$ は既知とす）

(b) $H_2: \theta_1 = \theta_2 = \cdots\cdots = \theta_k$ (但$A_1 = A_2 = \cdots\cdots = A_k$ なるもその値は未知)

(c) $H_3: \theta_1 = \theta_2 = \cdots\cdots \quad \theta_k$ (但A_i $i=1, 2, \cdots\cdots, k$は未知)

(B) A_i に関する仮説

(d) $H_4: A_1 = A_2 = \cdots\cdots = A_k$ (但 θ_i $i=1, 2, \cdots\cdots, k$は既知)

(e) $H_5: A_1 = A_2 = \cdots\cdots = A_k$ (但 $\theta_1 = \theta_2 = \cdots\cdots = \theta_k$ なるもその値は未知)

(f) $H_6: A_1 = A_2 = \cdots\cdots = A_k$ (但 θ_i, $i=1, 2, \cdots\cdots, k$は未知)

(g) $H_7: \theta_1 = \theta_2 = \cdots\cdots = \theta_k, \ A_1 = A_2 = \cdots\cdots = A_k$

次に各仮説に対する test statistics 及びその $p.d.f$ を挙げておく。

(A) に対してはH_1, H_2, H_3 のいづれも Bartlett L-test が応用出来る。

(a) $Z = \dfrac{2}{c} \left(R \log u - \sum_{i=1}^{k} r_i \log u_i \right)$ なる test statistics を使用せよ。

但 $Y_i = \sum_{j=1}^{r_i} (x_{ij} - A_i) + (n_i - r_i)(x_{ir_i} - A_i)$

$$U_i = \frac{Y_i}{r_i}, U = \frac{1}{R} \Sigma\, r_i u_i \quad R = \Sigma\, r_i, \quad C = 1 + \frac{1}{6(k-1)} \left(\Sigma \frac{1}{r_i} - \frac{1}{R} \right)$$

而るとき Z の $p.d.f$ は〔3〕より近似的 $\chi^2(k-1)$ -distribution 尙 r_i がすべて等しいときは power は多少おちるが $\min_i Y_i / \max_i Y_i$ を利用すればよい。そしてこの table はCadwell ·── Biometrika 1953 (筆者未見) ──によつて作られている。

(b) $Z = \dfrac{2}{c} \left[(R-1) \log u - \sum_{i \neq p} r_i \log u_i - (r_p - 1 \log u_p) \right]$ を使用する。

但 $Y' = \Sigma (x_{ij} - \hat{A}) + (n_i - r_i)(x_{ir_i} - \hat{A}) \quad \hat{A} = \min_i x_{i1}$

$\min_i x_{i1} = x_{p1}$ のとき $u_i = \dfrac{Y_i'}{r_i}$ $(i \neq p)$ $u_p = \dfrac{Y_p'}{r_p - 1}$

$$u = \frac{1}{R-1} \left\{ \Sigma\, r_i u_i + (r_p - 1) u_p \right\} \quad c = 1 + \frac{1}{6(k-1)} \left[\sum_{i \neq p} \frac{1}{r_i} + \frac{1}{r_p - 1} - \frac{1}{p-1} \right]$$

而るときこの Z は (a) と同じ分布法則に従う。

(c) $Z = \dfrac{2}{c} \left\{ (R-k) \log u - \Sigma (r_i - 1) \log u_i \right\}$ を使用すれば (a) と同じ分布をする

但 $\overset{*}{Y_i} = \Sigma (x_{ij} - x_{i1}) + (n_i - r_i)(x_{ir_i} - x_{i1})$

$$u_i = \overset{*}{Y_i} / (r_i - 1) \quad u = \frac{1}{R-k} \Sigma (r_i - 1) u_i \quad c = 1 + \frac{1}{6(k-1)} \left(\Sigma \frac{1}{r_i - 1} - \frac{1}{R-k} \right)$$

(d) $Z = 2 \sum_{i=1}^{k} \dfrac{n_i}{\theta_i} (x_{i1} - \hat{A}) \quad \hat{A} = \min_i x_{i1}$

を test statistcs として採用すれば此の分布は

$\chi^2(2h-2)$ ─ 分布になる

(e) $Z = \Sigma\, n_i (x_{i1} - \hat{A}) / \Sigma \overset{*}{Y_i}$

而るとき H_5 の仮説の下では $2 \Sigma \dfrac{n_i}{\theta} (x_{i1} - \hat{A})$: $\chi^2(2k-2)$ ─ 分布

又 $2\Sigma\dfrac{\overset{*}{Y_i}}{\theta}$: $\chi^2(2R-2)$ − 分布　で且両者は独立になることが証明出来るから Z の分布は由度 $(2k-2,\ 2k-2)$ の F 分布なることが示される。

(f) (g) に対しては large sample のときにのみ有効である検定方式しか求めることが出来なかつた。即ち H_6，H_7 に対して夫々

$$Z = m\prod_{i=1}^{k}\left(\frac{\overset{*}{Y_i}}{Y_i{}'}\right)^{ri},\qquad Z = \frac{m'\ \prod\ \overset{*}{Y_i}{}^{ri}}{\Sigma\left(\dfrac{Y_i{}'}{r_i}\right)^{R}}$$

$(m, m'$ は常数)

を用ぶれば $-2\log Z$ が r_i が相当大きいときに χ^2 − 分布することを利用してある程度目的が達せられる。

§ 3. 予備定理

§2 で説明した結果に直接必要な統計量の分布についての問題をこゝでとりあげる。今 $x_1\leqq x_2\leqq\cdots\cdots\leqq x_n$ を (1) からの ordered Sample とし truncated Sample を $x_1\leqq x_2\leqq\cdots\cdots\leqq x_n$ とする。而るときこの truncated Sample の同時的確率密度は

$$(2)\quad p(x_1,\cdots\cdots,x_r;\theta,A)=\begin{cases}\dfrac{n!}{(n-r)!}\dfrac{1}{\theta_r}\ exp\left\{-\dfrac{1}{\theta}\left[\sum_1^r(x_i-A)+(n_i-r_i)(x_r-A)\right]\right\} & A\leqq x_1\leqq\cdots\cdots\leqq x_r<\infty\\ O & \text{その他}\end{cases}$$

なることは明か。その時次の性質を証明する。

1. $1\leqq r\leqq n$ に対して $A\leqq c$ なる常数 c を考えると $x_1\geqq c$ なる条件の下で x_i-c の同時的確率密度は x_i-A のそれに等しい。

(証) $x_i-c=\overset{*}{x_i}$ とし，$\overset{*}{x_i}$ の同時的確率密度を $p(\overset{*}{x}\mid\overset{*}{x_1}\geqq c)$ とすれば $p(\overset{*}{x}\mid\overset{*}{x_1}\geqq c)=$

$$p(\overset{*}{x},\ \overset{*}{x_1}\geqq c)\ /\ p(\overset{*}{x_1}\geqq c)=\frac{n!}{(n-r)!}\frac{1}{\theta_r}\ exp\left\{-\frac{1}{\theta}\left[\Sigma\ \overset{*}{x_i}+(n-r)\overset{*}{x_r}+n(c-A\right]\right\}$$

$$/\,n\int_c^\infty\left(\int_{x1}^\infty\frac{1}{\theta}\ exp\left\{-\frac{1}{\theta}(x-A)\right\}dx\right)^{n-1}\frac{1}{\theta}\ exp\left\{-\frac{1}{\theta}(x_1-A)\right\}dx_1$$

$$=\frac{n!}{(n-r)!}\frac{1}{\theta^r}\ exp\left\{-\frac{1}{\theta}\left[\Sigma\ \overset{*}{x_i}+(n-r)\ \overset{*}{x_r}\right]\right\}$$

2. **1** に於ける常数 c を x_i と独立で (A,∞) で定義された確率変数 C におきかえても結果は変らない。

(証) C の $p.\ d.\ f.$ を $f(c)$ で表せば

$$P_r(X_1\geqq C)=\frac{n}{\theta}\int_A^\infty\int_c^\infty exp\left\{-\frac{n}{\theta}(X_1-A)\right\}f(c)\ dx_1\ dc$$

$$=\int_A^\infty exp\left\{-\frac{n}{\theta}(c-A)\right\}f(c)\ dc$$　より **1** と同様にして，証明可能。

3. $1 \leqq r \leqq n$ なる r に対して，

$W_j = (n-j+1)(x_j - x_{j-1})\ j=1,\ 2, \cdots\cdots, r$ 但 $x_o = A$

とすれば r 個の W_j は独立且 $\dfrac{2W_j}{\theta}$: $\chi^2(2)$ − 分布

更に $Y = \Sigma(\chi_i - A) + (n-r)(\chi_r - A)$ は $\chi^2(2r)$ − 分布

(証) $2W_j/\theta$ の特性函数を $\varphi(t_1, t_2, \cdots\cdots, t_r)$ とすれば

$$\varphi(t_1, \cdots\cdots t_r) = \frac{n!}{(n-r)!}\frac{1}{\theta_r}\int_a^\infty \cdots \int_a^\infty exp\left\{\frac{2i}{\theta}\Sigma t_j(n-+1)(x_j - x_{j-1}) - \frac{1}{\theta}\right.$$

$$\left.[\Sigma(x_j - A) + (n-r)(x_{rj} - A)]\right\} ax_i \cdots\cdots dx_r$$

この integral で $W_j = (n-j+1)(x_j - x_{j-1})$ なる変換をすれば

Jacobian は $\dfrac{(n-r)!}{n!}$ で $\sum_{j=1}^{r} W_j = \Sigma(x_i - A) + (n-r)(x_r - A)$

$$\therefore\ \varphi(t, \cdots\cdots, t_r) = \frac{1}{\theta_r}\int_o^\infty \cdots \int_o^\infty exp\left\{-\frac{1}{\theta}\sum_j(1-2itj)\,wj\right\} dw_1 \cdots\cdots dw_r$$

$$= \prod_{j=0}^{r}(1-2itj)^{-1}$$

一方 Wi のみの特性函数は $(1-2itj)^{-1}$ となり

r 個の Wj は 互に独立且夫々$x^2(2)$ − 分布に従う。

4. $Y' = \Sigma(X_i - X_1) + (n-r)(X_r - X_1)$ とおく。

$2Y'/\theta$ は X_1 と独立且 $x^2(2r-2)$ − 分布に従う。

(証) $Y' = \Sigma(X_j - X_1) + (n-r)(X_r - X_1) = \sum_{j=2}^{r} Wj$

3. より $W_1 = n(X_1 - A)$ は $Wj\ (j \neq 1)$ と独立従つて X^1 は $\sum_{j=2}^{r} Wj = Y'$ と独立になる。

更に Wj の独立性と各 々が $\chi^2(2)$ −分布に従うことから Y' は $\chi^2(2r-2)$ − 分布に従う。

5. (A, ∞) で定義され且 X_j と独立なる確率変数 X に対して $\overset{*}{Y} = \Sigma(X_j - X) + (n-r)(X_r - X)$

とおけば $X_1 \geqq X$ という条件の下で $2\overset{*}{Y}/\theta$ は $\chi^2(2r)$ − 分布に従う。

(証) 2. より簡単に出てくる。

6. $X_1, X_2, \cdots\cdots, X_n$ は互に独立且夫々次の確率密度をもつとする。

$$\frac{n_j}{\theta_j}\ exp\left\{-\frac{1}{\theta_j}(x_j - A)\right\}$$

而るとき $V = 2\sum_{j=1}^{k}\dfrac{n_j}{\theta_j}(X_j - \hat{A})\left(\text{但 } \hat{A} = \min_j x_j\right)$ は $\chi^2(2_k - 2)$ − 分布に従う。

(証) (Ⅰ) $\hat{A} = \min_j X_j = X_p$ な るとき $2\,\Sigma\,\dfrac{n_j}{\theta_j}(X_j - X_p) = V_p$ とす。

$X_j - A = Y_j$ とし同時的密度函数は，

$$\prod \frac{n_j}{\theta_j}\ exp\left\{-\Sigma\frac{n_j}{\theta_j}y_j\right\}$$

今 x_p が $\min_j xj$ なる probability を求めれば

$$P_r(x_1 > x_p,\ x_2 > x_p,\ \cdots\cdots,\ x_k > x_p) = P_r(y > y_p,\ \cdots\cdots,\ y_k > y_p)$$

$$= \prod \frac{n_j}{\theta_j}\int_0^\infty dy_p \int_{y_p}^\infty \cdots \int_{y_p}^\infty exp\left\{-\Sigma\frac{n_j}{\theta_j}y_j\right\}\prod_{j\neq p} dy_j \qquad = \frac{n_p}{\theta_p} \Big/ \sum_{1=j}^{k}\frac{n_j}{\theta_j}$$

次に $\min_j x_j = x_p$ という条件の下での V_p の条件付密度函数を $p(v_p \mid \min_j x_j = x_p)$ とすれば

$$p(v_p \mid min\ x_j = x_q) = \frac{p(u_p, minx_j\ x_j = x_p)}{p(minx_j = x_p)}$$

先ず $p(v_p,\ minx_j = x_p)$ を求めるため $x_p = minx_p$ のとき

V_p の特性函数を $\varphi(t)$ とすれば

$$\varphi(t) = E\left(e^{it2\Sigma\frac{n_j}{\theta_j}(y_j - y_p)}\right)$$

$$= \prod_j \frac{n_j}{\theta_j}\int_0^\infty dy_p\int_{y_p}^\infty\cdots\int_{y_p}^\infty exp\left\{-\Sigma\frac{n_j}{\theta_j}y_j + 2it\,\Sigma\frac{n_j}{\theta_j}(y_j - y_p)\right\}dy_1\cdots\cdots dy_k$$

$$= \prod_j \frac{n_j}{\theta_j}\int_0^\infty dy_p\int_{y_p}^\infty\cdots\int_{y_p}^\infty exp\left\{-(1-2it)\,\Sigma\frac{n_j}{\theta_j}y_j - y_p\left[\frac{n_p}{\theta_p} + 2it\sum_{j\neq p}\frac{n_j}{\theta_j}\right]\right\}\prod_{j\neq p}dy_j$$

$$= \frac{n_p}{\theta_p}\,\frac{1}{(1-2it)^{k-1}}\int_0^\infty exp\left\{-y_p\,\Sigma\frac{n_j}{\theta_j}\right\}dyp$$

$$= \frac{\frac{n_p}{\theta_p}}{\Sigma\frac{n_j}{\theta_j}}\,(1-2it)^{-(k-1)}$$

$$\therefore\ p(v_p,\ min\ x_j = x_p) = \frac{\frac{n_p}{\theta_p}}{\Sigma\frac{n_j}{\theta_j}}\,f_{2k-2}(v_p)$$

$f_{2k-2}(v_p)$ は $\chi^2(2k-2)$ － 分布の密度函数

$$\therefore\ p(v_p \mid min\ x_j = x_p) = f_{2k-2}(v_p)$$

(II) 次に V の確率分布を求めるために

$$P_r(V < v) = \sum_{p=1}^{k} P_r(V_p < v,\ min\ x_j = x_p)$$

$$= \Sigma\left(\frac{n_p}{\theta_p} \Big/ \Sigma\frac{n_j}{\theta_j}\right)F_{2k-2}(v) \qquad \text{但 } F_{2k-2}(v) : \chi^2(2k-2) - \text{分布の分布函数}$$

$$= F_{2k-2}(v)$$

即ち $V = 2\,\Sigma\frac{n_j}{\theta_j}(X_j - \hat{A})$ は $\chi^2(2k-2)$ － 分布に従う。

§ 4. 検定方式への解析

§ 3. の予備定理を基にして $H_1,\ H_3,\ H_4,\ H_5$ に対する解析を行う。他も大体同様である。

(1) なる $p.\ d.\ f$ をもつ k 母集団からの truncated Samples を

$X_{i1} \leqq X_{i2} \leqq \cdots\cdots \leqq X_{iri}\quad i = 1, 2, \cdots\cdots, k$ とし，尤度函数

$h(x_{i1},\ x_{i2},\ \cdots\cdots,\ x_{iri}\ ;\ \theta_i, A_i, i = 1, 2, \cdots\cdots, k)$ ──今後 $h(x;\theta, A)$ で表す──は

(3) $h(x;\theta,A) = \prod \frac{n_i\,!}{(n_i-r_i)\,!}\frac{1}{\theta_i^{r_i}} exp\left\{-\frac{1}{\theta_i}[\sum_j(x_{ij}-A_i)+(n_i-r_i)(x_{ir_i}-A_i)]\right\}$

1. H_1 に対する検定方式

$o<\theta_1, \theta_2, \cdots\cdots, \theta_k<\infty$ の張る k 次元母数空間を Ω とし，Ω に於いて $\max k(x;\theta,A)$ なる如き θ_i の estimates を $\hat{\theta}_i$ とすれば $\frac{\partial h}{\partial\theta_i} = o$ より

$$\hat{\theta}_i = \frac{1}{r_i}\left\{\Sigma(x_{ij}-A_i)+(n_i-r_i)(x_{ir_i}-A_i)\right\} = \frac{Y_i}{r_i}$$

次に $\theta_1=\theta_2=\cdots\cdots=\theta_k$ なる Ω の部分空間 ω に於いて $h(x;\theta,A)$ を maximum にする θ のestimate を $\hat{\theta}$ で表せば

$$\hat{\theta} = \frac{1}{R}\sum_i[\sum_j(x_{ij}-A_i)+(n_i-r_i)(x_{ir_i}-A_i)] = \frac{1}{R}\sum_i Y_i$$

故に尤度比 $\lambda = \frac{\max\limits_{\omega} h(x;\theta,A)}{\max\limits_{\Omega} h(x;\theta,A)} = \left(\frac{Y_1}{r_1}\right)^{r_1}\cdots\cdots\left(\frac{Yk}{rk}\right)^{rk}\Big/\left(\frac{1}{R}\Sigma Y_i\right)^R$

依つて H_1 なる仮説の下で λ の分布を求めれば

$\frac{Y_i}{r_i} = u_i \quad \frac{1}{R}\Sigma r_i u_i = u$ 　とおくとき

〔3〕から　$\chi^2 = \frac{2}{C}\left(R\log u - \Sigma r_i\log u_i\right)$

は r_i が比較的小さくも $\chi^2(k-1)$ − 布分をすると見做してもよい。

2. H_3 に対する検定方式

$\Omega:(o<\theta_1, \theta_2, \cdots\cdots,\theta_k<\infty, o<A_1,\cdots A_k<\infty)$ に於ける

A_i, θ_i の e stimate は夫々 $\hat{A}_i = \min\limits_j x_{ij}$

$$\hat{\theta}_i = \frac{1}{r_i}\left\{\Sigma(x_{ij}-x_{i1})+(n_i-r_i)(x_{ir_i}-x_{i1})\right\} = \frac{Y_i^*}{r_i}$$

$\omega:(o<\theta_1=\theta_2=\theta_k=\cdots\cdots=\theta_k<\infty, o<A_i<\infty$ なる Ω の部分空間)に於ける estimate は

$$\hat{A}_i = x_{i1},\ \hat{\theta} = \frac{1}{R}\sum_i[\Sigma(x_{ij}-x_{i1})+(n_i-r_i)(x_{ir_i}-x_{i1})] = \frac{\Sigma Y_i^*}{R}$$

依つて $\lambda = \left(\frac{Y_i^*}{r_i}\right)^{r1}\cdots\cdots\left(\frac{Y_k^*}{r_k}\right)^{rk}\Big/\left(\frac{1}{R}\Sigma Y_k^*\right)^R$

然るとき 1. と同様にして，

$\chi^2 = \frac{2}{c}\left\{(R-k)\log u - \Sigma(r_i-1)\log u_i\right\}$ は $\chi^2(k-1)$ − 分布に従う。

但 $\frac{Y_i^*}{r_i} = n_i,\quad \frac{1}{R-k}\Sigma r_i u_i = u$

3. H_4 に対する検定方式

$o<A_1, A_2, \cdots\cdots, A_k<\infty$ なる Ω に於ける A_i の estimate $\hat{A}_i = x_{i1}$ $A_1=A_2=\cdots\cdots A_k$ なる Ω の部分空間 ω に於いては

$\hat{A} = min\ x_{i1}$

$$\therefore \quad \lambda = exp\ \Sigma\ \frac{n_i}{\theta_i}\ (x_{i1} - \hat{A})$$

而して，§ 3. 6. に於いて $2\ \Sigma\ \frac{n_i}{\theta_i}\ (x_{i1} - \hat{A})$ は$\chi^2\,(2k-2)$ － 分布に従うことが証明されているから検定可能．

4. H_5 に対する検定方式

同様にしてλを求めれば

$$\lambda = \left(\frac{\Sigma\ n_i\,(x_{i1} - \hat{A})}{\Sigma\ \overset{*}{Y_i}} \right)^R$$

而してH_5 なる仮説の下では

$$2\ \Sigma\ \frac{n}{\theta}\ (x_{i1} - \hat{A})\ :\ \chi^2\,(2\,k - 2) - \text{分布}$$

$$\frac{2}{\theta}\ \Sigma\ \overset{*}{Y_i}\ :\ \chi^2\ \Big(\Sigma\,(2\,r_i - 2) \Big) - \text{分布}$$

更に § 3.4. よりx_{i1} は $\Sigma\ \overset{*}{Y_i}$ と独立であるから

$Z = \Sigma\ n_i(x_{i1} - \hat{A})\ /\ \Sigma\ \overset{*}{Y_i}$ は自由度 $(2k-2, 2R-2k)$ なるF分布に従う。

他の仮説に対する検定の解析も略同様である。そしていづれも既製の χ^2 又は F−table を利用して検定を行うことが出来る。

文　　献

〔1〕 B. Epstein and M. Sobel
Some Theo rems Relevant to Life -Testing from an Eexponen tial Distribution
Ann. Math Statist Vol 2 5 (1954)

〔2〕 B. Epstein aud C.K.Tsao
Some Tests Based onOrdercd Observations from Two Exponential Population
Ann. Math Statist Vol 2 3 (1952)

〔3〕 統計科学研究会編
統計数値表