昭和 37 年度　（問題）

午 前 の 部

1. 確率論における中心極限定理（central limit theorem）を説明し，これがどのように統計の理論および実際に応用されるか説明せよ。

2. 次の文中の〔　　〕の中の(A)～(J)について適当な式を入れよ。

任意の時刻で瞬間的に起る同種類の偶然変化を受ける確率過程（例えば電話の申込み呼数）を考え，変化が起れば過程の値は変化の起った時刻で一単位だけ順次飛躍してゆくとする。即ち時刻 t において $X_t = n$ であるならば，時間区間（ t，$t+h$ ）で 1 変化即ち次の値 $n+1$ への推移が起る確率は，$\lambda h + o(h)$ に等しく（ $\lambda > 0$ ），さらに（ t，$t+h$ ）間に 2 変化（即ち $n+2$ への推移）またはそれ以上の変化の起る確率は $o(h)$ に等しいとする。今時刻 t において $X_t = n$ である確率を $P_n(t)$ で表わすとこの仮定より

$$P_0(t+h) = P_0(t)(1-\lambda h) + o(h)$$

$$P_n(t+h) = P_n(t)〔(A)〕+ P_{n-1}(t)〔(B)〕+〔(C)〕,\quad n = 1, 2, 3, \cdots\cdots$$

が出る。$h \to 0$ ならしめて

(1)　$P_0'(t) = -\lambda P_0(t)$

(2)　$P_n'(t) = 〔(D)〕\ n = 1, 2, 3, \cdots$

を得る。初期条件として　$X_0 = 0$ とすると

$$P_n(0) = \begin{cases} 1 \ (n = 0) \\ 0 \ (n \neq 0) \end{cases}$$

である。(1)を解いて $P_0(t) = Ce^{-\lambda t}$ を得るが $P_0(0) = 1$ であるから $C = 〔(E)〕$ でなければならない。次に(2)を解くために $P_n(t) = e^{-\lambda t} U_n(t)$ とおき(2)に代入すれば

$$U_n'(t) = 〔(F)〕$$

が出る。$n > 0$ に対して $U_n(0) = P_n(0) = 0$ であるから

$$U_n(t) = \lambda \int_0^t U_{n-1}(s)\,ds$$

となる。しかし

$$U_0(t) = e^{\lambda t} P_0(t) = e^{\lambda t}〔(G)〕=〔(H)〕$$

であるから帰納的に

〔問〕

$$U_n(t) = 〔(I)〕$$

となる。従って

$$P_n(t) = 〔(J)〕$$

となるが，これは平均値がλtなるポワソン分布である。

午 後 の 部

1. 自由度nのt－分布の確率密度関数は

$$f_n(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}} \quad (-\infty < t < \infty)$$

で与えられる。

nが2より大きいとき，次のことを示せ。

(1) $E(t) = 0$

(2) $V(t) = \dfrac{n}{n-2}$

(3) t－分布を使う実例を一つあげ，簡単に説明せよ。

2. (1) 標本平均の分散は，比例層化抽出法によるものの方が，単純任意抽出法によるものより小であることを示せ。また，この場合，層別の効果を大きくするためには，どうすればよいかを述べよ。

(2) 母集団の大きさをN，母集団の変動係数$\left(\dfrac{\text{標準偏差}}{\text{平均値}}\right)$をcとするとき，単純任意抽出法において，推定値の変動係数c′を与えると，標本の大きさnは，次の式で決定できることを示せ。

$$n = \frac{N\dfrac{c^2}{c'^2}}{\dfrac{c^2}{c'^2} + N}$$

昭和 37 年度　（解答）

午 前 の 部

1. 中心極限定理は，二項分布の極限定理として知られる Laplace の定理の自然な拡張である。ある条件を満足する独立確率変数列の極限の分布が，正規分布であるという主張がこの定理の要点である。 この条件には，いろいろの形があるが，Lindeberg の条件が最も有名である。Lindeberg の条件とは，$\{X_i\}$ を与えられた分布函数 $\{F_i\}$ をもつ独立確率変数列とするとき，次の条件を満足するものをいう。

1° $\forall X_i$ は有限な分散をもつ。　　　$V(X_i)=\sigma_i^2<\infty$

2° $\forall X_i$ は有限な平均値をもつ。　　$E(X_i)=\mu_i<\infty$

3° 一般性を失うことなしに，$E(X_i)=0$ とするとき，

$\forall \varepsilon>0$ に対して

$$\frac{1}{\sum_{k=1}^{n} V(X_k)} \sum_{k=1}^{n} \int_{-\varepsilon\sqrt{\sum_{k=1}^{n} V(X_k)}}^{+\varepsilon\sqrt{\sum_{k=1}^{n} V(X_k)}} x^2\, dF_k(x) \to 1 \quad (n\to\infty)$$

が成立する。

中心極限定理は『Lindeberg の条件を満足する独立確率変数列について

$$\frac{\sum_{k=1}^{n} X_k}{\sqrt{\sum_{k=1}^{n} V(X_k)}}$$

の分布函数が，$N(0,1)$ に法則収束する』ということができる。

このように，中心極限定理は，正規分布との関係を明らかにしているために，統計の理論的展開の際に問題を正規分布に還元して議論できる基礎を与えているとみてよい。二項分布はもとより，Poisson 分布などを正規化してその性質を論ずることができる。

統計の応用の面からいえば，Lindeberg の条件をもつと強めたもので十分である。つまり，『$\{X_i\}$ がすべて，分散が有限な同一の分布に従う独立確率変数』とおきかえられる。この条件が，Lindeberg の条件を満足していることはいうまでもない。この条件を標本という立場から考えてみれば，母集団の分布が分散 σ^2 ，平均 μ をもっていることがわ

かれば，標本平均 $X = \dfrac{\sum_{k=1}^{n} X_k}{n}$ は，$N(\mu, \dfrac{\sigma^2}{n})$ に法則収束することが，中心極限定理によって保障されているのだから，検定，推定等に正規分布に関する結果を十分に利用することができる。

2. (A) $1-\lambda h$ (B) λh (C) $0(h)$ (D) $\lambda P_{n-1}(t) - \lambda P_n(t)$
   (E) 1 (F) $\lambda U_{n-1}(t)$ (G) $e^{-\lambda t}$ (H) 1 (I) $(\lambda t)^n / n!$
   (J) $\dfrac{(\lambda t)^n e^{-\lambda t}}{n!}$

午 後 の 部

1. (1) $f(t)$ が偶函数なること，またその$-\infty$から$+\infty$までの積分が収束することから明らか。$\left(f_n(t) = f_n(-t),\ f_n(t) \leqq \dfrac{A}{1+Bt^2},\ A>0,\ B>0 \right)$

   (2) $V(t) = \int_{-\infty}^{\infty} t^2 f_n(t)\, dt$

$$= \frac{2\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} t^2 \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt$$

$v = \dfrac{t^2}{n+t^2}$ と置換すれば，

$$= \frac{n\,\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)} \int_0^{1} v^{\frac{1}{2}} (1-v)^{\frac{n-2}{2}-1}\, dv$$

$$= \frac{n\Gamma(\frac{n+1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{n}{2})} \quad \frac{\Gamma(\frac{3}{2})\,\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n+1}{2})} = \frac{n\cdot\frac{1}{2}\Gamma(\frac{1}{2})\cdot\ \Gamma(\frac{n-2}{2})}{\Gamma(\frac{1}{2})\cdot\frac{n-2}{2}\ \ \Gamma(\frac{n}{2})}$$

$$= \frac{n}{n-2}$$

2. (1) 標本抽出法のうちの二方法，単純任意抽出法と比例層化抽出法の標本平均の分散の大小を比較するのであるが，単純任意抽出法の分散の公式は，既知とする。母集団の大きさをN，抽出数をn，母集団の分散をσ^2 とすれば，単純任意抽出法での標本平均の分散 V(rand) は，

$$V(\mathrm{rand}) = \frac{N-n}{Nn}\,\sigma^2$$ で与えられる。

次に，母集団を，kケに層化し，その各層の大きさをN_1，……，N_kとすれば，

(N_1 +……+ N_k = N) 比例層化抽出法では，各層から，$n_i = \frac{N_i}{N}n$ を抽出することになる。この場合の標本平均の分散 V(prop)は，各層の分散をσ_i^2 とすれば，

$$V(\mathrm{prop}) = \frac{N-n}{Nn}\sum_{i=1}^{k}\frac{N_i}{N}\sigma_i^2$$ で与えられる。

$$V(\mathrm{rand})-V(\mathrm{prop}) = \frac{N-n}{Nn}\sigma^2 - \frac{N-n}{Nn}\sum_{i=1}^{k}\frac{N_i}{N}\sigma_i^2$$

$$= \frac{N-n}{Nn}\left(\sigma^2 - \sum_{i=1}^{k}\frac{N_i}{N}\sigma_i^2\right) \quad \cdots\cdots\cdots\cdots\cdots ①$$

ここで，$\left.\begin{array}{l}\widetilde{X}：母集団平均\\ \widetilde{X}_i：第 i 層の平均\end{array}\right\}$ とすれば，$\sigma^2 = \frac{1}{N}\sum_{\alpha=1}^{N}(X_\alpha - \widetilde{X})^2$

$$= \sum_{i=1}^{k}\frac{N_i}{N}\sigma_i^2 + \sum_{i=1}^{k}\frac{N_i}{N}(\widetilde{X}_i - \widetilde{X})^2$$

であるから，①式は，

$$V(\mathrm{rand}) - V(\mathrm{prop}) = \frac{N-n}{Nn} \cdot \sum_{i=1}^{k} \frac{N_i}{N} (\tilde{X}_i - \tilde{X})^2 \geqq 0$$

よって，$V(\mathrm{rand}) \geqq V(\mathrm{prop})$ である。

上式より，層別の効果は，

$$\frac{N-n}{Nn} \sum_{i=1}^{k} \frac{N_i}{N} (\tilde{X}_i - \tilde{X})^2$$ で与えられるから

この値を大にすればよい。このことは，各層ごとの平均がなるべく異る値をとるようにすればよい。ということは，同質のものを層化するのが効果的だといえる。

(2) 推定値の変動係数を c' とすれば，c と c' の間の関係は

$$V(\mathrm{rand}) = \frac{N-n}{Nn} \sigma^2 \text{ より } c'^2 = \frac{N-n}{Nn} c^2 \text{ である。}$$

これより，n をとけば $n = \dfrac{N \dfrac{c^2}{c'^2}}{\dfrac{c^2}{c'^2} + N}$ となる。