昭和54年度（問　題）

1.　不良率がおよそ10%であることが知られている製品について，標本不良率と真の不良率との差が5%以下になることを確率0.99で保証するためには，何個のサンプルを必要とするか。ただし，標準正規分布の両側ε点を$u(\frac{\varepsilon}{2})$で表すとき，$u(0.005)=2.575$とする。

2.　A，B 2班の人に同じ問題でテストを行ったところ，A班は平均73点，標準偏差14点で，B班は平均76点，標準偏差18点であった。A班の人数を20人，B班の人数を30人とするとき，B班の方が優秀といえるか。必要あれば次の数字を用いよ。

$$\frac{1}{\sqrt{2\pi}}\int_0^{1.96} e^{-\frac{x^2}{2}}dx = 0.475\ ,\quad t_{19}(0.05)=2.09\ ,\quad t_{29}(0.05)=2.05\ ,\quad t_{48}(0.05)=2.01\ ,$$

$$F^{29}_{19}(0.025)=2.24\ ,\quad F^{29}_{19}(0.975)=0.446$$

3.　ある州のA候補の支持率pを調査するため，n人を対象に調査したところ，k人の人が支持していることがわかった。nはかなり大きいとして，pを次の手順に従って区間推定せよ。

(ア)　積率母関数を使って，$U=(\hat{p}-p)\Big/\sqrt{\dfrac{p(1-p)}{n}}$は正規分布$N(0,1)$に従うことを示せ。ここに$\hat{p}=\dfrac{k}{n}$とする。

(イ)　正規分布$N(0,1)$の両側ε点を$u(\frac{\varepsilon}{2})$として，支持率pを区間推定せよ。

4.　一様分布$U(0,1)$に従う独立な確率変数X_1，X_2，$\cdots$，X_nの順序統計量を$X_{(1)}$，$X_{(2)}$，$\cdots$，$X_{(n)}$とする。ただし，$X_{(1)}\leqq X_{(2)}\leqq\cdots\leqq X_{(n)}$。

$n\to\infty$のとき，$P[nX_{(1)}>t]$，$P[nX_{(2)}>t]$の極限値を求めよ。ここに$t>0$とする。

5.　$\{X_i\}(i=1,2,\cdots)$を，共通の密度関数$f(x)=\alpha e^{-\alpha x}(x\geqq0,\ \alpha>0)$を持つ独立な確率変数とし，$S_0=0$，$S_n=X_1+X_2+\cdots+X_n(n=1,2,\cdots)$とする。ここで，確率変数$N(t)$を次のように定義する：

$N(t)$は$S_k\leqq t$となるようなインデックス$k\geqq1$の個数

すなわち，

$$[N(t)=n]=[S_n\leqq t,\quad S_{n+1}>t]$$

このとき，

(ア)　S_nの分布関数$G_n(x)$を求めよ。

(イ)　$N(t)$はいかなる分布に従うか。

昭和54年度（解答例）

1. サンプル数をn，真の不良率をpとすると，標本不良率$\hat{p}$は近似値に正規分布

$N\left(p,\ \dfrac{p(1-p)}{n}\right)$ に従うことが知られている。

ゆえに，統計量

$$U=\frac{\hat{p}-p}{\sqrt{\dfrac{p(1-p)}{n}}}$$

は正規分布 $N(0,\ 1)$ に従う。

つまり，確率 $1-\varepsilon$ で

$$|\hat{p}-p|<u\left(\frac{\varepsilon}{2}\right)\sqrt{\frac{p(1-p)}{n}}$$

が成り立つ。

よって，$\varepsilon=0.01$，$p=0.1$として

$$u(0.005)\sqrt{\frac{0.1\times0.9}{n}}\leqq0.05$$

をnについて解けばよい。

$$n\geqq\left(\frac{2.575}{0.05}\right)^2\times0.09\fallingdotseq238.7$$

およそ240以上のサンプルが必要である。

2. A班，B班の得点が，おのおの正規分布 $N(\mu_x,\sigma_x{}^2)$，$N(\mu_y,\ \sigma_y{}^2)$ に従っているとして検定を行う。

(1) まず，等分散性の検定を行う。

帰無仮説 $H_0:\sigma_x{}^2=\sigma_y{}^2$，対立仮説 $H_1:\sigma_x{}^2\neq\sigma_y{}^2$ につき，不偏分散比 F_0を求める。

$$F_0=\left(\frac{30\times18^2}{29}\right)\Big/\left(\frac{20\times14^2}{19}\right)=1.625$$

一方，$F_{19}^{29}(0.025)=2.24$であるから，有意水準５％で分散は等しいとみなして

よい。

(2)　(1)で分散は等しいとみなしてよいから，帰無仮説 $H_0 : \mu_x = \mu_y$，対立仮説 H_1: $\mu_x \neq \mu_y$ を t-検定すると，

$$|t| = \frac{|73-76|}{\sqrt{\dfrac{20\times14^2+30\times18^2}{20+30-2}\left(\dfrac{1}{20}+\dfrac{1}{30}\right)}} = 0.61649 < t_{48}(0.05)$$

よって，有意水準5％で差がある（すなわち，B班の方が優秀である）とはいえない。

3.　(ア)　X_iを，調査対象者$i\,(1\leqq i\leqq n)$がA候補を支持していれば1，支持していなければ0なる値をとる確率変数とすると，$\hat{p}=\dfrac{1}{n}\sum_{i=1}^{n}X_i$，ここで，$Y_i = X_i - p$ とおくと，

$$E(Y_i) = 0,\quad E(Y_i{}^2) = V(X_i) = p(1-p)$$

よって，Y_iの積率母関数$\phi(\theta)$は，θが小さいとき

$$\phi(\theta) = 1+\frac{1}{2}p(1-p)\theta^2+0(\theta^3)$$

と表わすことができる。従って，$U = \sum_{i=1}^{n}Y_i \Big/ \sqrt{np(1-p)}$ の積率母関数 $\phi_n(\theta)$ は，

$$\phi_n(\theta) = \left\{\phi\left(\theta\Big/\sqrt{np(1-p)}\right)\right\}^n$$

$$= \left\{1+\frac{\theta^2}{2n}+0(n^{-\frac{3}{2}})\right\}^n$$

$\dfrac{\theta^2}{2n}+0(n^{-\frac{3}{2}})=x$ は，$n\to\infty$のとき $\to 0$であるから

$$\phi_n(\theta) = \left\{(1+x)^{\frac{1}{x}}\right\}^{\frac{\theta^2}{2}+0(n^{-\frac{1}{2}})} \longrightarrow e^{\frac{\theta^2}{2}}\ (n\to\infty)$$

ここで $e^{\frac{\theta^2}{2}}$ は $N(0,\ 1)$の積率母関数であるから求める結果が得られた。

(イ)　(ア)より，確率 $1-\varepsilon$ でUは $\left(-u\left(\frac{\varepsilon}{2}\right),\ u\left(\frac{\varepsilon}{2}\right)\right)$ に入る。

これを書きかえると

$$\hat{p}-u\left(\frac{\varepsilon}{2}\right)\sqrt{\frac{p(1-p)}{n}}<p<\hat{p}+u\left(\frac{\varepsilon}{2}\right)\sqrt{\frac{p(1-p)}{n}}$$

である。求める区間は平方根の中の p を $\hat{p}$ に置きかえたものである。

4. $P\left[nX_{(1)}>t\right]=P\left[X_{(1)}>\frac{t}{n}\right]$

であるが，$X_{(1)}$の定義から

$$=P\left[X_1>\frac{t}{n},\ \cdots,\ X_n>\frac{t}{n}\right]$$

となる。ここで$\left\{X_i\right\}_{i=1,\cdots,n}$は独立で $U(0,1)$に従うから

$$=\prod_{i=1}^{n}P\left[X_i>\frac{t}{n}\right]$$

$$=\left(1-\frac{t}{n}\right)^n$$

ここで $n\to\infty$とすると

$$\longrightarrow e^{-t}$$

となる。

次に，$P\left[nX_{(2)}>t\right]$ の $n\to\infty$ の極限値も同様に求めると

$$P\left[nX_{(2)}>t\right]=P\left[X_{(2)}>\frac{t}{n}\right]$$

$$=P\left[X_{(1)}>\frac{t}{n}\right]+P\left[X_{(1)}\leqq\frac{t}{n},\ X_{(2)}>\frac{t}{n}\right]$$

$$=\left(1-\frac{t}{n}\right)^n+\binom{n}{1}\frac{t}{n}\left(1-\frac{t}{n}\right)^{n-1}$$

$$\longrightarrow e^{-t}+te^{-t}$$

となる。

5. (ア) S_n の密度関数 $g_n(x)$ を求める。

$$g_1(x)=\alpha e^{-\alpha x}\qquad (x\geqq 0)$$

は明らかで，次に

$$g_k(x) = \alpha \frac{(\alpha x)^{k-1}}{(k-1)!} e^{-\alpha x} \qquad (x \geqq 0)$$

と仮定して，$g_{k+1}(x)$ を求めると

$$g_{k+1}(x) = \int_0^x g_k(x-t) g_1(t)\, dt$$

$$= \int_0^x \frac{\alpha^k}{(k-1)!}(x-t)^{k-1} e^{-\alpha(x-t)} \cdot \alpha e^{-\alpha t}\, dt$$

$$= \frac{\alpha^{k+1}}{(k-1)!} e^{-\alpha x} \int_0^x (x-t)^{k-1}\, dt$$

$$= -\frac{\alpha^{k+1}}{k!} e^{-\alpha x} (x-t)^k \Big|_0^x$$

$$= \alpha \frac{(\alpha x)^k}{k!} e^{-\alpha x} \qquad (x \geqq 0)$$

より帰納法で $g_n(x)$ が求められた。

$$\therefore G_n(x) = \int_0^x g_n(t)\, dt$$

$$= \frac{\alpha^n}{(n-1)!} \int_0^x t^{n-1} e^{-\alpha t}\, dt$$

$$= \frac{\alpha^n}{(n-1)!} \left\{ -\frac{1}{\alpha} t^{n-1} e^{-\alpha t} \Big|_0^x + \frac{n-1}{\alpha} \int_0^x t^{n-2} e^{-\alpha t}\, dt \right\}$$

$$= -\frac{(\alpha x)^{n-1}}{(n-1)!} e^{-\alpha x} + \frac{\alpha^{n-1}}{(n-2)!} \int_0^x t^{n-2} e^{-\alpha t}\, dt$$

と順次積分を繰り返して t の次数を落とすと

$$G_n(x) = 1 - e^{-\alpha x} \left(1 + \frac{\alpha x}{1!} + \frac{(\alpha x)^2}{2!} + \cdots + \frac{(\alpha x)^{n-1}}{(n-1)!} \right)$$

が得られる。

(イ) $P[N(t) = n] = P[S_n \leqq t,\ S_{n+1} > t]$

$$= G_n(t) - G_{n+1}(t)$$

$$= e^{-\alpha t}\frac{(\alpha t)^n}{n!}$$

となり，ポアソン分布に従うことがわかる。