基礎統計(水 1)・補足シケプリ　文責・$U-yan$

このプリントではこの前アップした既製シケプリ 1～3 で手薄になっているところを中心に補足説明していきます．順番は教科書通りでいきます．

【第 1 章～第 7 章】
この部分に関しては教科書や既製シケプリを見て下ればそれほど問題ないと思います．そこで 2 つだけ忘れやすい注意すべき事柄を．

○教科書 P114 ポアソンの小数の法則

「$np\to\lambda, n\to\infty, p\to 0$ となる極限に於いて，任意の整数 x に対し ${}_nC_x p^x(1-p)^{1-x}\to\dfrac{e^{-\lambda}\lambda^x}{x!}$ 」．

この法則はべらぼうに n が大きいけど，p がかなり小さい二項分布について確率を求めたいときに使います．例えば，$p=0.002$ の稀な現象が起こるかどうかを $n=1000$ 回観測するとき，この現象が 2 回起こる確率は $f(2)={}_{1000}C_2(0.002)^2\cdot(0.998)^{998}$ ですがこんなの計算できません(試験場で電卓の「＝」を 1000 回以上押すことになります)．そこで，$np=2$ ですから，ポアソンの小数の法則を用いて

$f(2)\fallingdotseq\dfrac{e^{-2}2^2}{2!}\fallingdotseq 0.271$ と求めることになります．理論ではなく実用上有益な法則です．(そもそも統計は実学の色彩が濃いですし)

○教科書 P148　7.4 節 独立な確率変数の和
この節は後半に出てくる標本分布を理解するのに実は重要な部分です．この節で述べているのは，

「n 個の確率変数 X_1，X_2，$\cdots$，X_n について，期待値の和に関する次の関係が成り立つ．

(ⅰ)　$\boldsymbol{E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n)}$　(和の期待値＝期待値の和)

さらに，n 個の確率変数 X_1，X_2，$\cdots$，X_n が独立ならば，分散に関して次が成り立つ．

(ⅱ)　$\boldsymbol{V(X_1+X_2+\cdots+X_n)=V(X_1)+V(X_2)+\cdots+V(X_n)}$　(和の分散＝分散の和)　」

という，二つの式です．(ⅱ)の「独立ならば」という条件は結構重要で，(ⅰ)は常に成立するのですが，(ⅱ)は常には成り立ちません．注意しましょう．証明はそれほど難しくないですし確かシケプリ 1 にあったので割愛します．

で，この関係式がなぜ重要かというと，先ほども述べましたように，これはある確率分布 X から n 個の値を取り出してきたとき，つまり n 個の**確率変数** X_1，X_2，$\cdots$，X_n をとってきた時に重要な役割を果たすからです．確率分布 X(母分布)の平均を μ，分散を σ^2 とします．
このとき，X_1，X_2，$\cdots$，X_n はそれぞれ独立で，平均 μ，分散 σ^2 の確率分布に従うと考えられますから 確率変数の和 $S_n=X_1+X_2+\cdots+X_n$ について，先程の関係式(ⅰ)(ⅱ)を用いれば

$E(S_n)=E(X_1+X_2+\cdots+X_n)=n\mu$，$V(S_n)=V(X_1+X_2+\cdots+X_n)=n\sigma^2$　が成り立ちます．

さらに，確率変数の平均 $\bar{X}=\dfrac{X_1+X_2+\cdots+X_n}{n}$ について，期待値や分散の性質を用いれば，

$$E(\bar{X})=E\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)=\frac{1}{n}E\left(X_1+X_2+\cdots+X_n\right)=\mu \quad ,$$

$$V(\bar{X})=V\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)=\frac{1}{n^2}V(X_1+X_2+\cdots+X_n)=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$$ が成り立ちます．

この二つは 8 章の中心極限定理と通ずるものがありますし，9 章・10 章でたまに出てきます．

※ この場合の X_1，X_2，・・・，X_n は定数ではなく確率**変数**であることに注意です．

【第 8 章】

○中心極限定理(*Central Limit Theorem*; CLT)

「母集団分布が何であっても，そこからとりだした n 個の標本の和 $X_1+X_2+\cdots+X_n$ は，n が十分大きいとき，近似的に正規分布に従う．則ち，母集団平均を μ，母集団分散を σ^2 とすれば，

$S_n=X_1+X_2+\cdots+X_n$ は $\mathrm{N}(n\mu,\ n\sigma^2)$ に，$\bar{X}=\dfrac{X_1+X_2+\cdots+X_n}{n}$ は $\mathrm{N}(\mu,\ \dfrac{\sigma^2}{n})$ に近似的に従う．」

この定理は実用的に大変重要で，具体的に問題を解くときに使うことも結構あります．母集団分布の形を仮定しないでもその和・平均が正規分布に漸近していくというのは考えてみると結構凄いです．

<u>例題</u>「細工のされていない(各目が出る確率が均等の)さいころを n 回振るとき，1 の目が出る回数を r 回とする． $\left|\dfrac{r}{n}-\dfrac{1}{6}\right|\leq\dfrac{1}{100}$ となる確率が 0.99 以上になるような最小の n を求めよ．」

<u>[解答の流れ]</u> さいころを i 回目に投げたとき，一の目が出たら $X_i=1$，そうでなければ $X_i=0$ となる確率変数を用いると，$r=X_1+X_2+\cdots+X_n$ である．各々の X_i は二項分布 $Bi(1,\ \dfrac{1}{6})$ に従い，その平均は $1\cdot\dfrac{1}{6}=\dfrac{1}{6}$，分散は $\dfrac{1}{6}\cdot\left(1-\dfrac{1}{6}\right)=\dfrac{5}{36}$ である．ゆえに，中心極限定理より(母集団分布を $Bi(1,\ \dfrac{1}{6})$ とみる)，n が十分大きいとき，$r=X_1+X_2+\cdots+X_n$ は近似的に正規分布 $\mathrm{N}(\dfrac{n}{6},\ \dfrac{5n}{36})$ に従う．ゆえにそれを n で割った相対度数 $\dfrac{r}{n}$ は正規分布 $\mathrm{N}(\dfrac{1}{6},\ \dfrac{5}{36n})$ に近似的に従う．あとはこれを標準化して，

区間 $\left|\dfrac{r}{n}-\dfrac{1}{6}\right|\leq\dfrac{1}{100}$ が 99%信頼区間に入るようにするだけ．□

ちなみに上の定理の記述は大分おおざっぱで，正確に記すならば以下のようになります．

「$n\to\infty$ のとき，$P\left(a\leq\dfrac{X_1+X_2+\cdots+X_n-n\mu}{\sqrt{n}\sigma}\leq b\right)\to\displaystyle\int_a^b\frac{1}{2\pi}e^{-x^2/2}dx$ 」．

【第 9 章～第 12 章】

この範囲に関しては既製シケプリ 1~3 に詳しく書かれていますので，そちらを参照してくだされば問題ないと思います．しかしモーメント法・最尤法による推定の話の扱いが小さいので少し解説します．

○推定

そもそも推定とは，ある母集団から得られた有限個の標本から，その母集団分布を推測することです．現実問題では経験的に母集団分布の種類（正規分布やパレート分布，などの○○分布）が分かっていて，あとはその**母数**（その分布の形を決める値；例えば正規分布では平均 μ と分散 σ^2 の 2 つ）を推定することが多いです．以下，得られた標本 X_1，X_2，･･･，X_n (今回は具体的な定数です)から母数を推定する方法を 2 つ紹介します．

①モーメント法

まず，モーメントの定義から述べますと，確率変数 X に対して，
原点の周りの k 次のモーメントを $E(X^k)$ で定義します．そしてモーメント法をごく大雑把に言えば，

(母集団分布から計算される k 次のモーメント)＝(標本から計算される k 次のモーメント)

とおいて，方程式を解いて母数を求める，というものです．

具体的にいえば, ある m 個の母数 $\theta_1,\theta_2,\cdots,\theta_m$ で決定される母集団分布 $F(\theta_1,\theta_2,\cdots,\theta_m)$ があったとき，まず，定義通りに 1 次から m 次までのモーメントを計算します(分布の式が与えられていれば計算できます)．この式は，**未知数である母数 $\theta_1,\theta_2,\cdots,\theta_m$ を含んだ式**です．
そして次に，n 個の標本 X_1，X_2，･･･，X_n から標本モーメントをこれまた 1 次から m 次まで計算します．具体的に n 個得られた値に関しては 期待値＝平均 ですので(もしくは一様分布と考えてもよい)，

(k 次の標本モーメント)＝$\dfrac{1}{n}\sum_{i=1}^{n} X_i^{\,k}$ です．（これは具体的な値）

最後に，(母集団分布から計算される k 次のモーメント)＝(標本から計算される k 次のモーメント)
とおいてやれば, これは m 個の未知数 $\theta_1,\theta_2,\cdots,\theta_m$ についての m 個(1 次～m 次)の方程式ですから, 各々の θ の具体的値を求める事ができるわけで，それをそれぞれ $\theta_1,\theta_2,\cdots,\theta_m$ の推定値とします．

例えば X が正規分布 N(μ，σ^2) に従うと分かっている時，μ，σ^2 の推定をしてみます．今，標本として具体的値 X_1，X_2，･･･，X_n が得られたとします．まず，母集団分布から 1 次と 2 次のモーメントを計算しますが(母数が 2 つです！)，正規分布の性質として我々は既に X の平均が μ で分散が σ^2 であることを知っています．すなわち，$E(X)=\mu$，$V(X)=\sigma^2$ です．さらに，教科書 P97 の(5・28)より

$V(X)=E(X^2)-\left(E(X)\right)^2$ ですから，$E(X^2)=V(X)+\left(E(X)\right)^2=\sigma^2+\mu^2$ となります．

ゆえに連立方程式 $\begin{cases} E(X)=\dfrac{1}{n}\sum_{i=1}^{n}X_i \\ E(X^2)=\dfrac{1}{n}\sum_{i=1}^{n}X_i^{\,2} \end{cases}$ ，すなわち $\begin{cases} \mu=\dfrac{1}{n}\sum_{i=1}^{n}X_i \\ \sigma^2+\mu^2=\dfrac{1}{n}\sum_{i=1}^{n}X_i^{\,2} \end{cases}$ を解いて，

推定値 $\begin{cases} \hat{\mu} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i = \overline{X} \\ \hat{\sigma}^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2 = S_X{}^2 \end{cases}$ を得ます.

○最尤法(さいゆうほう)

最尤法とは読んで字のごとく最も尤度(もっともらしさ；*likelihood*)を大きくするものを推定値とするものです．標本のような結果が得られる最ももっともらしい分布を推定分布として採用する訳です.

具体的にはまず，ある未知母数 $\theta_1, \theta_2, \cdots, \theta_m$ を含む母集団分布 $F(\theta_1, \theta_2, \cdots, \theta_m)$ があったとき，その**確率**分布 $f(X,\ \theta_1, \theta_2, \cdots, \theta_m)$ を用いて，具体的に得られた標本値 X_1，X_2，･･･，X_n に対して尤度関数 L を次のように定義します.

$$L(\theta_1, \theta_2, \cdots, \theta_m) \equiv \prod_{i=1}^{n} f(X_i, \theta_1, \theta_2, \cdots, \theta_m)$$

$f(X_i, \theta_1, \theta_2, \cdots, \theta_m)$ は標本値 X_i が得られる確率ですから，尤度関数は母集団から n 個の値をとってきた時に，**<u>X_1，X_2，･･･，X_n のような標本が得られる確率</u>**を表します．そしてこの尤度関数 L を最大にするような（$\theta_1, \theta_2, \cdots, \theta_m$）を推定値とするのが，最尤法です．最大値をとるような $\theta_1, \theta_2, \cdots, \theta_m$ を求めるのには，偏微分 $\dfrac{\partial L}{\partial \theta_k} = 0 \quad (k = 1, 2, \cdots, m)$ を用います.

なお，積のままでは計算しにくいので対数尤度 $\log\left(L(\theta_1, \theta_2, \cdots, \theta_m)\right) = \displaystyle\sum_{i=1}^{n} \log\left(f(X_i, \theta_1, \theta_2, \cdots, \theta_m)\right)$ を考えることも多いです.

例えば，再び，X が正規分布 N（μ，σ^2）に従うと分かっていて，標本 X_1，X_2，･･･，X_n が得られた時に，平均 μ と分散 σ^2 の推定をしてみます．この場合，X_i という値が得られる確率，つまり確率分布は

$f(X_i, \mu, \sigma^2) = \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(-\dfrac{(X_i - \mu)^2}{2\sigma^2}\right)$ ですから，尤度関数は

$L(\mu, \sigma^2) = \displaystyle\prod_{i=1}^{n} \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left(-\dfrac{(X_i - \mu)^2}{2\sigma^2}\right)$ となります．積のままでは計算しづらいので対数尤度

$$\log\left(L(\mu, \sigma^2)\right) = \log\left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(X_i - \mu)^2}{2\sigma^2}\right)\right) = \sum_{i=1}^{n} \log\left(\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(X_i - \mu)^2}{2\sigma^2}\right)\right)$$
$$= n \log \frac{1}{\sqrt{2\pi}\sigma} - \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{2\sigma^2} = -n \log \sqrt{2\pi}\sigma - \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{2\sigma^2}$$

を使って（σ^2 を推定するので σ^2 で微分して） $\begin{cases} \dfrac{\partial \log\left(L(\mu, \sigma^2)\right)}{\partial \mu} = \displaystyle\sum_{i=1}^{n} \dfrac{X_i - \mu}{\sigma^2} = 0 \\ \dfrac{\partial \log\left(L(\mu, \sigma^2)\right)}{\partial \sigma^2} = -\dfrac{n}{2\sigma^2} + \displaystyle\sum_{i=1}^{n} \dfrac{(X_i - \mu)^2}{2\sigma^4} = 0 \end{cases}$

を解くと，モーメント法と同じ結果
$$\begin{cases} \hat{\mu} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i = \overline{X} \\ \hat{\sigma}^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} (X_i - \overline{X})^2 = S_X{}^2 \end{cases}$$
を得ます.

※この場合はモーメント法と最尤法の結果が一致しましたが，一般には 2 つの方法での推定値は異なります.

※「偏微分が 0」と「尤度関数が最大」は同値ではないですが，数学的な吟味は省略しています.

【第 13 章】

テストには出ないので割愛です．本当はここも統計の華なのですが授業が 13 回しかないので辿り着けなかったようです.

(以上)