平成 27 年 7 月 31 日（金） 第５限

確率・統計学 （春学期） 期末試験

【問１】,【問２】を**１枚目の解答用紙**、【問３】,【問４】を**２枚目の解答用紙**に解答せよ。
解答用紙に**学籍番号**と**名前**の記入を忘れないようにすること。
単純ミスがあっても途中点を与えられるように，考え方の筋道が分かるように解答すること。

必要に応じて，裏面の表を用いよ。

【問１】(30) 以下のように関数 $f(x)$ を定義する。

$$f(x) = \begin{cases} c\,|x| & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

以下の問いに答えよ。

(1) $f(x)$ が確率密度となるように定数 c を求めよ。そのうえで，平均 μ と分散 σ^2 を求めよ。

(2) $f(x)$ の特性関数 $F(k)$ を求めよ。ここで，特性関数 $F(k)$ は $F(k) = E\left[e^{ikX}\right]$ で定義される。

(3) (2) の $F(k)$ を用いて平均 μ と分散 σ^2 を求め，(1) の結果と一致することを示せ。

【問２】(30) 以下の問いに答えよ.

(1) ある事象 A が確率 p で起こるとする。これを n 回試行したとき，A が x 回起こるときの 2 項分布の確率関数 $f(x)$ を書き下せ。ここで，x は 0 から n の整数である。そのうえで，$f(x)$ が規格化されていることを示せ。さらに，二項分布における平均 μ と分散 σ^2 を $f(x)$ を用いて求めよ。

(2) 平均が μ のポアソン分布の確率関数 $p(x)$ を書き下せ。ここで，x はゼロ以上の整数とする。そのうえで，$p(x)$ が規格化されていることを示せ。さらに，ポアソン分布における分散 σ^2 を $p(x)$ を用いて求めよ。

(3) 中心極限定理を述べよ。

【問３】(10) 母平均が μ，母分散が σ^2 の正規母集団から大きさ n の標本を抽出した。標本平均 $\bar{X}$ と標本分散 S^2 の定義式を書いたうえで，期待値 $E\left[\bar{X}\right]$ および $E\left[S^2\right]$ を求めよ。

【問４】(30) 以下の問いに答えよ.

(1) 標本の大きさが $n = 2$ の標本平均 $\bar{X}$ が母平均 μ の有効不偏推定量であることを示せ。ここで，有効不偏推定量とは，分散の期待値が最小となる不偏推定量のことである。

(2) 母分散 $\sigma^2 = 9$ の正規母集団から大きさ n の標本を抽出して，母平均 μ の信頼水準 99%の信頼区間を求めたい。その区間の幅を 1 以下にするには，標本の大きさ n をいくつ以上にしないといけないか。

(3) 工場の機械を調整して，平均 1000 mℓ，標準偏差 10 mℓ の正規分布となるように設定したとする。ペットボトルを 100 本無作為に取って水の量を測ったら，その平均が 1003 mℓ であった。この機械は正しく調整されているといえるか，帰無仮説を明示したうえで有意水準 2 % で検定せよ。

確率・統計学（春学期） 期末試験 [解答例]

【問１】

(1) 正規化条件から,

$$\int_{-\infty}^{\infty} dx\, f(x) = \int_{-1}^{1} dx\, c\,|x| = 2c\int_{0}^{1} dx\, x = c = 1$$

よって，$c = 1$。

平均は,

$$\mu = E\left[X\right] = \int_{-1}^{1} dx\, x\,|x| = 0$$

ここで，被積分関数は奇関数であることを用いた。同様に，分散は,

$$\sigma^2 = E\left[X^2\right] - \mu^2 = \int_{-1}^{1} dx\, x^2\,|x| = 2\int_{0}^{1} dx\, x^3 = \frac{1}{2}$$

(2) 特性関数 $F(k)$ は

$$F(k) = E\left[e^{ikX}\right] = \int_{-1}^{1} dx\ |x|\, e^{ikx} = \int_{0}^{1} dx\, 2x\cos(kx) = \frac{2\left(k\sin k + \cos k - 1\right)}{k^2}$$

あるいは,

$$F(k) = \frac{2 - \left(e^{ik} + e^{-ik}\right) + ik\left(e^{ik} - e^{-ik}\right)}{(ik)^2}$$

(3) $k = 0$ 近傍でテイラー展開すれば，いずれも

$$F(k) = 1 - \frac{k^2}{4} + \frac{k^4}{72} + \cdots$$

となるから,

$$\mu = \left.\frac{\partial F(k)}{\partial(ik)}\right|_{k=0} = 0, \qquad \sigma^2 = E\left[X^2\right] - \mu^2 = \left.\frac{\partial^2 F(k)}{\partial(ik)^2}\right|_{k=0} = \frac{1}{2}$$

となって，(1) の結果と一致する。

【問２】

(1) 二項分布で与えられるから,

$$f(x) = \frac{n!}{x!\,(n-x)!}p^x(1-p)^{n-x} = \frac{n!}{x!\,(n-x)!}p^x q^{n-x}$$

二項定理から,

$$\sum_{x=0}^{n} f(x) = (p+q)^n = 1$$

平均は,

$$\mu = E\left[X\right] = \sum_{x=1}^{n} x f(x) = p\frac{\partial}{\partial p}\sum_{x=1}^{n}\frac{n!}{x!(n-x)!}p^x q^{n-x} = pn(p+q)^{n-1} = np$$

同様に,

$$\sigma^2 = E\left[X^2\right] - \mu^2 = p\frac{\partial}{\partial p}p\frac{\partial}{\partial p}\sum_{x=1}^{n}\frac{n!}{x!(n-x)!}p^x q^{n-x} - \mu^2 = np - np^2 = npq$$

(2) ポアソン分布で与えられるから，

$$p(x) = e^{-\mu}\frac{\mu^x}{x!}$$

また，

$$\sum_{x=0}^{\infty} p(x) = e^{-\mu}\sum_{x=0}^{\infty}\frac{\mu^x}{x!} = e^{-\mu}e^{\mu} = 1$$

分散は，

$$\sigma^2 = E\left[X^2\right] - \mu^2 = \sum_{x=0}^{\infty} x^2 p(x) - \mu^2 = e^{-\mu}\mu\frac{\partial}{\partial\mu}\mu\frac{\partial}{\partial\mu}\sum_{x=0}^{\infty}\frac{\mu^x}{x!} - \mu^2 = \mu + \mu^2 - \mu^2 = \mu$$

(3) n 個の独立な確率変数を $X_1, X_2, \cdots, X_n$ とする。このとき，$\bar{X} = (X_1 + X_2 + \cdots + X_n)/n$ で定義される確率変数 $\bar{X}$ の確率分布は，n が十分に大きいとき，正規分布 $N\left(\mu, \frac{\sigma^2}{n}\right)$ で近似できる。ここで，μ と σ^2 はそれぞれ X_i（$i = 1, \cdots n$）の平均と分散である。

【問３】大きさが n の標本平均は $\bar{X} = (X_1 + X_2 + \cdots + X_n)/n$ で定義される。
期待値を取る演算子 $E[...]$ は線形だから，

$$E\left[\bar{X}\right] = \frac{1}{n}\sum_{k=1}^{n} E\left[X_k\right] = \frac{1}{n}n\mu = \mu$$

標本分散の定義は，

$$S^2 = \frac{1}{n}\sum_{k=1}^{n}\left(X_k - \bar{X}\right)^2$$

期待値は，

$$\begin{aligned} E\left[S^2\right] &= E\left[\frac{1}{n}\sum_{k=1}^{n}\left(X_k - \bar{X}\right)^2\right] = \frac{1}{n}E\left[\sum_{k=1}^{n}\left\{(X_k - \mu)^2 + \left(\bar{X} - \mu\right)^2 - 2\left(X_k - \mu\right)\left(\bar{X} - \mu\right)\right\}\right] \\ &= \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n}\sigma^2 \end{aligned}$$

【問４】

(1) 不偏推定量を $\hat{\theta} = aX_1 + bX_2$ で定義する。
これが母平均 μ の不偏推定量であるためには，$E\left[\hat{\theta}\right] = \mu$ である必要がある。よって，

$$a + b = 1$$

さらに，分散を計算すると，

$$\begin{aligned} V\left[\hat{\theta}\right] &= E\left[\left(\hat{\theta} - \mu\right)^2\right] = E\left[a^2(X_1 - \mu)^2 + b^2(X_2 - \mu)^2 + 2ab\left(X_1 - \mu\right)\left(X_2 - \mu\right)\right] \\ &= \sigma^2\left(a^2 + b^2\right) = 2\sigma^2\left(a - \frac{1}{2}\right)^2 + \frac{\sigma^2}{2} \end{aligned}$$

よって，$a = b = \frac{1}{2}$ で最小となるから，有効不偏推定量 $\hat{\theta}$ は標本平均 $\bar{X}$ と一致する。

(2) 標準正規分布のもとでの信頼水準 99%の信頼区間は，数表から $[-2.576, 2.576]$ である。よって，信頼区間は，

$$-2.576 \leq \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq 2.576 \quad \Rightarrow \quad \bar{X} - 2.576\frac{3}{\sqrt{n}} \leq \mu \leq \bar{X} + 2.576\frac{3}{\sqrt{n}}$$

この区間が 1 以下であるためには，

$$2 \times 2.576\frac{3}{\sqrt{n}} \leq 1 \quad \Rightarrow \quad n \geq 238.9$$

$n \in \mathbb{N}$ だから，$n \geq 239$。

(3) 機械が正しく設定されているときの平均を帰無仮説に選ぶ。よって，

H_0 :　$\mu = 1000$ mℓ

平均からのずれが問題となるので，両側検定である。よって，対立仮説は

H_1 :　$\mu \neq 1000$ mℓ

標本の大きさは $n = 100$ で，母平均と母標準偏差はそれぞれ $\mu = 1000$, $\sigma = 10$ だから，有意水準 2 % の採択域は

$$-2.326 \leq Z = \frac{\bar{X}-1000}{1} \leq 2.326 \quad \Rightarrow \quad 997.7 \leq \bar{X} \leq 1002.3$$

$\bar{X} = 1003$ mℓ は棄却域に入るから，有意水準 2 % で帰無仮説は棄却され，対立仮説が採択される。つまり，機械は正しく設定されていないといえる。