統計学入門－第２回レポート
水尾　2009/12/22

出題範囲－第２章

注意事項－解答レポートは原則手書きで作成すること。表やグラフも全て直筆(フリーハンド or 定規)で描くこと。レポート用紙のサイズ・種類は問わない。提出レポートは解答部分のみで良い。
問題の解き方も説明すべき問は、解答までの道筋をきちんと述べよ。また、計算は結果のみならず途中の計算式も分かる様に示せ。ただし、問題文中に計算結果のみ示す旨の指示がある場合は計算結果のみで良い。

提出期限－2010年1月12日(火)17：00

第１問(離散型確率の標本空間と独立性)

目の数	1	2	3	4	5	6
確率	1/10	1/10	3/10	1/5	1/5	1/10

右表に示した確率をもつ変形した特殊さいころ３回投げで、一回目のさいころの目の数がl(l=1～6)，二回目のさいころの目の数がm(m=1～6)，三回目のさいころの目の数がn(n=1～6)となる根元事象を$e_{l,m,n}$と記すことにする。また、一回目，二回目，三回目のさいころの目の数を表す確率変数をそれぞれX,Y,Zとする。

(1)　一回目と二回目のさいころの目の数が等しくかつ３回のさいころの目の数の総合計が13になる事象Aを根元事象の集合A＝{根元事象,根元事象,…,根元事象}の形式で示せ。

(2)　３回の目の数の総合計が偶数になる事象Bの確率P(B)を求めよ。

(3)　確率変数X＋Y＋Zが４になる確率$P_{X+Y+Z}(X+Y+Z=4)$を求めよ。

ヒント(2)　問題集のヒントを参考にせよ。

第２問(離散型確率の標本空間と確率変数)

大きさ８の(抽象的な)標本空間$\Omega=\{e_1, e_2, \cdots, e_8\}$を考える。事象$A=\{e_1, e_3, e_5, e_7\}$，$B=\{e_2, e_4, e_6\}$，$C=\{e_1, e_8\}$とし、それぞれの事象の確率はP(A)＝1/3，P(B)＝1/2，P(C)＝1/3であるとする。

(1)　事象$D=\{e_3, e_5, e_7\}$の確率P(D)を求めよ。

(2)　この標本空間上で次のような(抽象的な)確率変数Xを考える。確率変数Xの値は、事象Bが起きたときには１、事象Dが起きたときには－1、事象$(B\cup D)^c$が起きたときには０をとるものとする。このとき確率変数Xと確率変数|X|の確率分布を表で示せ。また、確率変数Xと|X|の平均値μ_X, $\mu_{|X|}$、分散σ_X^2, $\sigma_{|X|}^2$を求めよ。

第３問(特性値の公式)

確率変数	平均値(g)	分散(g^2)
X	－1.0	1
Y	1.0	100
Z	10.0	10000

離散型確率変数X,Y,Z(単位グラムg)の３つを考える。確率変数YとZは独立であるが､確率変数XとZの独立性は不明である。また、各確率変数の平均値と分散は右表のように与えられているとする。

(1)　以下の各確率変数の平均値と分散を求めよ。ただし、以下の確率変数の特性値のなかには、問題に与えられた条件を用いるのみでは求めることが出来ないものも作為的に含まれている。その場合には｢解答不可能｣と解答せよ。
①4X+6　②Y^2＋4Y　③X＋2Z＋2　④Y＋2Z＋2

(2)　確率変数X＋bZの平均値が区間[－10.0,200.0)に入るための未定係数bの条件を求めよ。

(3)　確率変数bY＋cZの平均値が0.0、分散が100であるとき、未定係数b,cのちの組(2組)を求めよ。

ヒント(1)②　$\sigma_Y^2 = E((Y-\mu_Y)^2) = E(Y^2 - 2\mu_Y Y + \mu_Y^2) = E(Y^2) - 2\mu_Y E(Y) + \mu_Y^2$

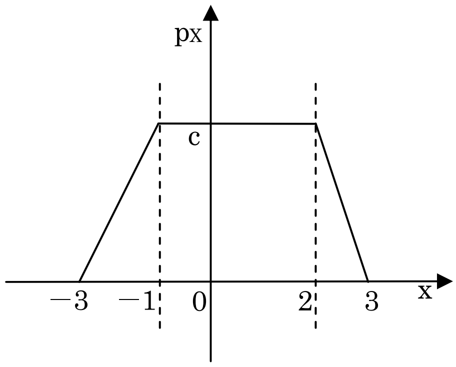
第４問(独立性)

(普通の)さいころ100回投げで、m回目(m＝1～100)のさいころの目の数を表す確率変数をX_m(m＝1～100)としたとき、確率変数$\frac{1}{100}\sum_{m=1}^{100} X_m$の平均値と分散を求めよ。

<u>考察</u>　確率変数$\frac{1}{100}\sum_{m=1}^{100} X_m$は 100 回の目の数の記述統計の意味でのデータの平均値を意味している。したがって、確率変数$\frac{1}{100}\sum_{m=1}^{100} X_m$の確率論の意味での平均値と分散は、『記述統計の意味でのデータの平均値』の確率論の意味での平均値・分散である。この考え方が第 3 章推測統計の基礎となる。

第 5 問(連続型確率変数の確率密度関数)

ある連続型確率変数 X の確率分布を表す確率密度関数 $p_X(x)$が右図のようであるとする。



(1)　グラフ中の c の値を決定せよ。

(2)　確率変数 2.5X の確率密度関数 $p_{2.5X}(x)$のグラフを示せ。

第 6 問(連続型確率変数の確率密度関数)

ある連続型確率変数 X の確率分布が区間[1,3)の一様分布であるとする。また、区間事象 A=[0.0,2.4),B=(−∞,2.5),C=[1.5,2.0)とする。事象 D が以下のように与えられたとき、事象 D を区間事象の和事象の形式に書き直せ。また、確率変数 X の値が事象 D に入る確率 $P_X(D)$も求めよ。

(1)　$D=A^c$のとき

(2)　D=B−C のとき

(3)　D=(B∩A)−C のとき

(4)　$D=(B\cap C^c)\cup A$のとき

第 7 問(正規分布表と標準化)

次の各値を正規分布表を用いて求めよ。

(1)　X〜N(0,1)のとき下流確率 $P_X((-\infty, -0.25))$の値

(2)　X〜N(0,1)のとき上流確率 $P_X([2.13,\ \infty))$の値

(3)　X〜N(0,1)のとき区間確率 $P_X([-0.30,\ a))=0.9266$ となるようなおよその a の値

(4)　X〜N(2,100)のとき区間確率 $P_X([-3.0,\ 28.1))$の値

(5)　X〜N(10,10000)のとき区間確率 $P_X([-20,\ a))=0.5561$ となるようなおよその a の値

第 8 問(正規分布)

確率変数 X〜N(μ,σ²)のとき、$P_X([\mu-\sigma,\mu+\sigma))$,$P_X([\mu-2\sigma,\mu+2\sigma))$,$P_X([\mu-3\sigma,\mu+3\sigma))$をそれぞれ求めよ。

<u>考察</u>　この問題の結果は、「正規分布では、平均値(=分布の中央)の周りに幅 2σ(=2 標準偏差分),4σ,6σの区間を考えたときのその区間の確率」がいくつであるかということを教えている。

第 9 問(標準化と特性値の公式)

確率変数 X〜N(2.0,100),Y〜N(20.0,1)で X と Y は独立であるとするとき以下の各問に答えよ。

(1)　確率変数 X を標準正規分布を確率分布とする確率変数に変換するための標準化の式を示せ。

(2)　(1)の標準化で、X の区間[−10,100)は Z のどのような区間に変換されるかを示せ。

(3)　確率 $P_X([-10,100))$を求めよ。

(4)　確率変数 X+100Y の平均値と分散を求めよ。

(5)　確率変数 X+bY の平均値が区間[−10.0,200.0)に入るための未定係数 b の条件を求めよ。

(6)　確率変数 X+bY の分散が区間[400,900)に入るための未定係数 b の条件を求めよ。

(7)　未定係数 b を実数値全体の範囲で動かした場合に確率変数 X+bY の分散のとりうる最小値を求めよ。

(8)　確率変数 bY+cZ の平均値が 0.0、分散が 100 であるとき、未定係数 b,c の値の組(2 組)を求めよ。

第10問(独立性)

　ある特殊なウイルスTに侵された場合、侵された人間の体温が36.95℃以上かつ血糖値が170.0mg/dl以上であるときは、直ちに(潜伏期間ほとんどなしに)そのウイルスは急速に増殖し重い肺炎を誘発して患者はほぼ確実に亡くなるが、体温または血糖値がそうではないときは、このウイルスの増殖は白血球によっていとも簡単に抑え込まれ完全に死滅するとする。このウイルスに侵されたある人が肺炎を併発して死に至る確率を求めよ。ただし、人間の体温の平均値は36.50℃で分散は0.0400(℃)^2(←この2つの数値は適当に作問した)、血糖値の平均値は98.2mg/dlで分散は400.00(mg/dl)^2(←この2つの数値も適当に作問した)であり、体温と血糖値は確率論的に独立(←この独立性も実際にはあり得ない)な関係にあるほぼ正規分布の状態(←この正規分布性も実際にはあり得ない)であるとして計算せよ。

ヒント　講義で配布した統計数値表よりもっと大きな統計数値表が必要なときは、図書館やインターネットで入手できる。