統計学Ⅰ　2016 年度春学期　問題と解答
以下の設問の解答は解答用紙（別紙）を半分に折り左右２段組みにしたうえで**問題順**に記しなさい。＊印の問題では答えに至る途中過程も記しなさい

数表
標準正規分布 Z について P(Z＜1.645)=0.95, P(Z＜1.96)=0.975

1 標本統計量に関する以下の問いに答えよ。(20)
(1)標本平均 $\bar{x}$ や標本分散 s^2 が持つ共通の望ましい性質を何というか。
(2) $\bar{x}$ や s^2 が持つ共通の欠点は何か。
(3) $\bar{x}$ の代わりに(2)の欠点を補う統計量として何が利用されるか。一つ挙げよ。
(4)変動係数の定義とその利点を簡潔に述べよ。
解答　各 5 点
(1)不偏性　4 点(普遍性 etc 誤記、不偏性と効率性)、0 点(BLUE 等)
(2)外れ値の影響を受けやすい
(3)中央値または刈り込み平均　0 点(変動係数、歪度、尖度　過去問の弊害？)
(4) $s/\bar{x}$, 水準に差のある場合の散らばりの比較に利用できる　定義は言葉で説明しても可だが $\bar{x}/s$ は×　3 点(定義 or 利点のみ正解)、利点として散らばりの尺度だけは×

2　ある学生 i が近視である $x_i=1$ か否か $x_i=0$ がベルヌイ分布 $f(x_i)=\theta^{x_i}(1-\theta)^{1-x_i}$ に従うとき以下の問いに答えよ。(25)
(1*)確率変数 x_i の平均 $E(x_i)$ と分散 $Var(x_i)$ を定義から導出せよ。
(2)無作為に選んだ n 人の学生のうちの近視の学生の数 $y=\Sigma x_i$ はどんな確率分布に従うか。
(3)無作為とはどのような状況かを簡潔に説明しなさい。
(4)n を増やしていくと y はどのような確率分布(平均と分散も答えよ)に従うか。
解答
(1) $E(x)=1\times\theta+0\times(1-\theta)=\theta$, $Var(x)=(1-\theta)^2\times\theta+(0-\theta)^2\times(1-\theta)=\theta(1-\theta)[(1-\theta)+\theta]=\theta(1-\theta)$　各 5 点、4 点($Var(x)=E(x^2)-\mu^2=$...分散公式は定義ではない)　0 点($E(x)=\theta$ や $Var(x)=\theta(1-\theta)$ の答えだけ、ベルヌイ分布は離散確率変数なのに $\int xf(x)dx$ など連続確率変数の誤った定義は×)
(2)2 項分布　5 点　0 点(正規分布、ベルヌイ分布)
(3)互いに独立である(標本選択が他の標本の影響を受けない)　5 点　0 点(ランダム、偏りがない)　ランダムは無作為の英訳で答えにはならない、標本は「無作為かつ偏りなく」選ぶ必要があるが、無作為と偏りのないは異なる基準です
(4)平均 $n\theta$ 、分散 $n\theta(1-\theta)$ の正規分布　5 点　3 点(正規分布のみ)　2 点(標準正規分布、0 点にしたいけど、この回答が多かった。なぜ Σx の平均が 0 に近づくのか？ありえない)

3　(2 の続き)無作為に選んだ n 人の学生の近視の割合 $p=y/n$ について以下の問いに答えなさい。(25)
(1*)近視の割合 p の平均 $E(p)$ と分散 $Var(p)$ をモーメント公式より θ と n の式にしなさい。
(2*)n=100 の調査で p=0.5 であったとき、母数 θ の 95%信頼区間を求めなさい。
(3*)「学生の近視率 θ は 0.8 である」という仮説を有意水準 5%で検定しなさい。
解答
(1) $E(y/n)=1/nE(y)=1/n(n\theta)=\theta$ 、 $Var(y/n)=1/n^2Var(y)=1/n^2\{n\theta(1-\theta)\}=\theta(1-\theta)/n$　各 5 点　*がついているので答えだけは 0 点　モーメント公式の演算が出来るか否かが問われている
(2) $0.5-1.96\sqrt{0.5\cdot0.5}/\sqrt{100}=0.402<\theta<0.598=0.5+1.96\sqrt{0.5\cdot0.5}/\sqrt{100}$　5 点　4 点(1.96 を 2 で近似)　3 点(計算ミス)　0 点(1.645 で計算)
(3) $H_0:\theta=0.8$, $H_1:\theta<0.8$ とする。帰無仮説 $H_0:\theta=0.8$ が真であるとすると、 $p\sim N(0.8,(0.04)^2=$

$\sqrt{0.8\cdot 0.2}/\sqrt{100}$）となるが、p=0.5 を標準化すると z=$\dfrac{p-\theta}{\sqrt{\theta(1-\theta)}/\sqrt{n}}$=(0.5−0.8)/0.04=−7.5 <−1.645 なので、帰無仮説 H_0 は棄却される。ゆえに有意水準 5%で 0.8 未満と言える。

10 点満点　各ステップで誤りがあると 3 点減点した。多い誤りは①仮説設定の誤り（H_0: θ=0.5、H_0: μ<0.8 など）②標準化の誤り（分散を p=0.5 で計算して z=(0.5-0.8)/0.05=−6）③臨界値の誤り（1.96）　④結論の誤り　など。検定は重要なのに出来が悪かった。秋学期は検定が出来ないと単位が取れなくなりますよ。

4　都心のあるバス停では、待ち時間 X が 0～5 分の一様分布に従うという。（20）
(1)確率変数 X の確率密度関数 f(X)を求めなさい。
(2*)分布関数 F(X)を定義から導出しなさい。
(3*)確率変数 X の平均 E(X)と分散 Var(X)を求めなさい。
解答
(1)f(X)=1/5　5 点　1/6 という人がいたが、0.1.2…,6 のような離散確率変数ではない　密度関数と書いてあるから連続確率変数

(2)F(X)= $\int_0^X f(x)dx = \frac{1}{5}\left[x\right]_0^X = \frac{1}{5}X$　5 点　定義を間違えて F(X)=1 とする人が散見された

(3)E(X)= $\int_0^5 xf(x)dx = \frac{1}{5}\left[\frac{1}{2}x^2\right]_0^5 = \frac{1}{10}25$ =2.5、Var(X)=E(x^2)−$\left(\frac{5}{2}\right)^2 = \frac{100-75}{12} = \frac{25}{12}$

$\int_0^5 x^2 f(x)dx = \frac{1}{5}\left[\frac{1}{3}x^3\right]_0^5 = \frac{1}{15}125 = \frac{25}{3}$　各 5 点

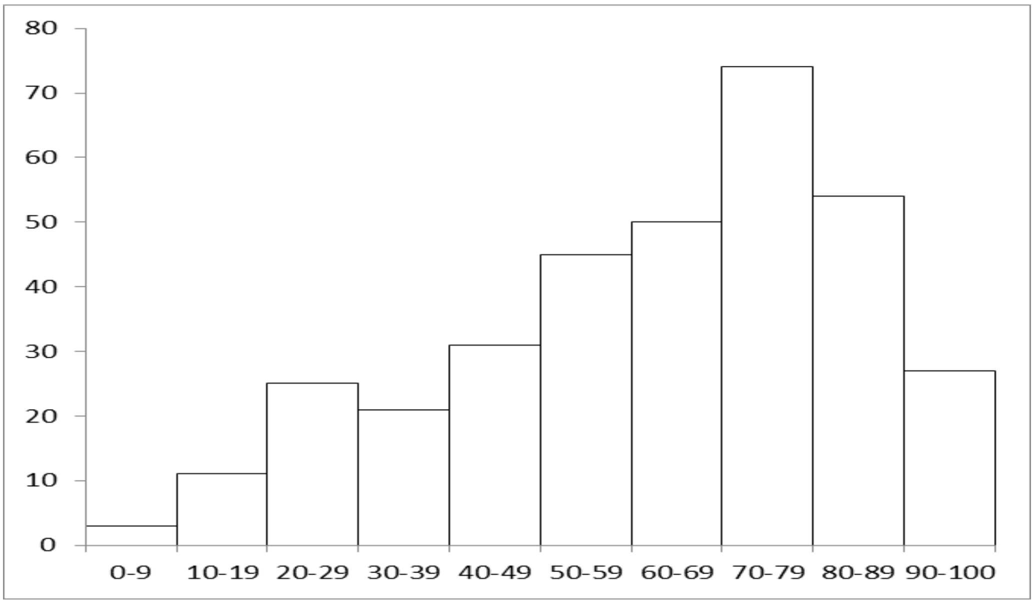
すべて計算ミスは 3 点とした

5　ある企業が作る部品の国別生産シェアが国内 $P(X_J)$=50%、アジア $P(X_A)$=30%、米国 $P(X_U)$=20%で、それぞれの不良品率が $P(Y|X_J)$=1%、$P(Y|X_A)$=3%、$P(Y|X_U)$=2%であるとき、以下の問いに答えよ。(10)
(1*)この会社の不良品率 P(Y)を求めなさい。
(2*)この会社の不良品はどこで作られた確率が最も高くなるか。
解答
(1) 全確率の公式より
$P(Y)=P(Y\cap X_J)+P(Y\cap X_A)+P(Y\cap X_E)$=0.5*0.01+0.3*0.03+0.2*0.02=18/1000　5 点
(2)ベイズの定理より
日本産 $P(X_J|Y)=P(Y\cap E_J)/P(Y)$=(5/1000)÷(18/1000)=5/18
アジア産 $P(X_A|Y)=P(Y\cap X_A)/P(Y)$=(9/1000)÷(18/1000)=9/18=0.5
米国産 $P(X_E|Y)=P(Y\cap X_E)/P(Y)$=(4/1000)÷(18/1000)=4/18
なのでアジア産である確率が最も高くなる。　5 点　*がついているから答えだけは減点

採点結果

受験者数　341 人　平均点　62 点　標準偏差 22 点　最高点 100 点　最低点 0 点

下記は期末試験(100 点満点)の度数分布です。



多くの学生が 70 点以上でしたが、50 点未満の人もやや多いです。

成績は「総点＝レポート乗数の平均値×期末試験得点」に基づき、A(総点≧80)　B(80>総点≧60)　C(60>総点≧40)　D(総点＜40)でつけます。総点＜50 を D としたら D の割合が高くなったので、条件を緩めました。