数理統計学　前回の解答　　第２回分

１）表が出る確率を $p=\dfrac{3}{4}$ としたとき、６回目に初めて表が出る確率。

幾何分布の形に合わせると

「５回連続して裏が出た後、６回目に初めて表が出る確率」

なので、$p=\frac{3}{4}, k=5$ を代入すればよい。よって確率は、

$$P(X=5)=\left(1-\frac{3}{4}\right)^5\times\frac{3}{4}=\frac{1}{1024}\times\frac{3}{4}=\frac{3}{4096}\left(=0.0073242\cdots=0.00732=7.32\times10^{-3}\right)$$

２）表が出る確率を $p=\dfrac{2}{3}$ としたとき、４回中３回表が出る確率。

$p=\frac{2}{3}, n=4, k=3$ の二項分布の確率を計算すればよい。

$$\begin{aligned}P(X=3) &= {}_4\mathrm{C}_3\times\left(\frac{2}{3}\right)^3\times\left(1-\frac{2}{3}\right)^{4-3}=\frac{2\times3\times4}{1\times2\times3}\times\left(\frac{2}{3}\right)^3\times\left(\frac{1}{3}\right)^1\\ &= 4\times\frac{8}{27}\times\frac{1}{3}=\frac{32}{81}\left(=0.39506\cdots=0.395\right)\end{aligned}$$

３）表が出る確率を $p=\dfrac{2}{5}$ としたとき、５回中２回裏が出る確率。

５回中２回裏が出るということは、５回中３回表が出ればよいので、
$p=\frac{2}{5}, n=5, k=3$ の二項分布の確率を計算すればよい。

$$\begin{aligned}P(X=3) &= {}_5\mathrm{C}_3\times\left(\frac{2}{5}\right)^3\times\left(1-\frac{2}{5}\right)^{5-3}=\frac{3\times4\times5}{1\times2\times3}\times\left(\frac{2}{5}\right)^3\times\left(\frac{3}{5}\right)^2\\ &= 10\times\frac{8}{125}\times\frac{9}{25}=\frac{144}{625}\left(=0.2304=0.230=2.30\times10^{-1}\right)\end{aligned}$$

《別解》裏の出る確率のまま、$p=1-\frac{2}{5}=\frac{3}{5}, n=5, k=2$ の二項分布の確率でも良い。

$$P(X=2)={}_5\mathrm{C}_2\times\left(\frac{3}{5}\right)^2\times\left(1-\frac{3}{5}\right)^{5-2}=\frac{4\times5}{1\times2}\times\frac{9}{25}\times\frac{8}{125}=\frac{144}{625}$$

《注意》何人かの学生が解答を「１－（裏が３回出る確率）」としているが、
これでは「裏が３回ではない確率」＝「表が１回以外になる確率」となる。

数理統計学において、確率の値は既約分数（これ以上約分できない分数）または有効数字３桁以上で答えることとする。既約分数でない場合や、有効数字が３桁に満たない場合は１点減点となるので注意。有効数字については裏面を参照のこと。

有効数字について

工学の世界では実際に何かを測定した場合、有効数字として記入された桁数までは信頼できる値であることを示しているため、有効数字を正確に取り扱うことは非常に重要である。本来であれば、途中計算の桁数や最終結果の桁の処理方法（いわゆる四捨五入や切捨て等）に関しても色々と細かい計算規則があるが、この授業ではごく一般的に用いられている有効桁数＋１桁目を四捨五入したものと考える。

１）有効数字の桁数について

有効数字の定義はいくつかあるが、ここでは簡単に先頭から連続した 0 を除いて、残った数字の数を有効数字とする。例えば

12.345 $\rightarrow$ 有効数字 5 桁
0.123 $\rightarrow$ 有効数字 3 桁 $\Leftarrow$ 左側にある“ 0. ”の部分は有効桁数として数えない
0.0012 $\rightarrow$ 有効数字 2 桁 $\Leftarrow$ 左側にある“ 0.00 ”の部分は有効桁数として数えない

のようになる。次に 1 万や 1 億などの大きな桁数の数字の場合で有効数字の桁数を本来の桁数よりも小さくする場合は、次のようにする。

１万を有効数字３桁 $\rightarrow$ 1.00×10^4 “ 10,000 ”と書くと有効数字 5 桁になる
７億２千万を有効数字２桁 $\rightarrow$ 7.2×10^8 “ 720,000,000 ”と書くと有効数字 9 桁になる

つまり、$a.bcd \times 10^n$ の形で書くと有効桁数が分かりやすい。

12.345 $\rightarrow$ 1.2345×10
0.123 $\rightarrow$ 1.23×10^{-1}
0.0012 $\rightarrow$ 1.2×10^{-3}

注意）１万を有効数字３桁書く場合、右側の「0」には意味があるので、
間違っても消してはならない。

２）演算による有効数字の変化

有効数字を考慮した場合四則演算によって有効数字の桁数が変わる場合があるので注意が必要である。特に有効数字の桁数が違うものの演算を行ったとき、ほとんどの場合において精度の悪いほうに揃う（つまり有効桁数が上がることはない！）。また同じような数の引き算を行うと、有効桁数が低くなるので注意が必要となる。

この授業においては、特に断りがなければ**有効数字３桁以上**で答えることとする（演算による有効数字の変化については特に厳しく見ない）。有効数字が良く分からない場合は、とりあえず電卓の結果をすべて書き表せば試験では正解とする。ちなみに、有効数字の桁数が足りない場合（注意：電卓の結果が「0.25」の場合の解答は「0.250」となる）、および分数で記入した場合に既約分数（完全に約分されている）で無い場合は減点の対象（通常－１点）となるので、注意してください。

数理統計学 確率変数と確率分布　　第 3 回

・確率分布と連続型確率の計算について

確率分布の定義は離散型、連続型を問わず、次のように x 以下になる確率で定義される。

$$F(x) = P(-\infty < X \leqq x) \qquad \left(F(-\infty) = 0, \quad F(\infty) = 1 \right)$$

また連続型の場合に確率分布が微分可能であるとき、その導関数

$$f(x) = \frac{d}{dx} F(x)$$

のことを X の確率密度関数と呼ぶ。このとき、連続型の確率は積分を使って

$$P(a < X \leqq b) = \int_a^b f(x) dx$$

で計算できる。この式から連続型において一点 x になる確率の値は 0 であることがわかる。さらに、$P(a < X \leqq b) = P(a \leqq X \leqq b) = P(a < X < b)$ である。

・確率分布の例（連続型の場合）

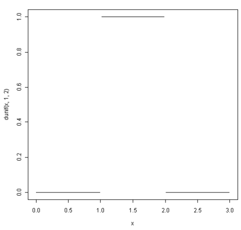
・一様分布 (パラメータは a, b)

ある区間 $[a, b]$ で一様に分布している。実際には確率密度関数が常に一定の分布。

確率密度関数 : $f(x) = \dfrac{1}{b-a} \quad (a \leqq x \leqq b)$

分布関数 : $F(x) = \dfrac{x-a}{b-a} \quad (a \leqq x \leqq b)$

平均 : $\mu = \dfrac{a+b}{2}$　　分散 ： $\sigma^2 = \dfrac{(b-a)^2}{12}$



例　区間 $[1, 4]$ の一様分布に従う確率変数 X が区間 $(2, 3]$ に入る確率

$$P(2 < X \leqq 3) = \int_2^3 \frac{1}{3}\, dx = \left[\frac{1}{3} x \right]_2^3 = \frac{2}{3} - \frac{1}{3} = \frac{1}{3} = 0.3333 \cdots = 0.333$$

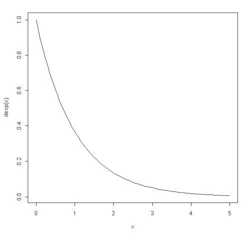
・指数分布 (パラメータは λ)

統計でよく用いられる基本的な分布の一つ。到着時間間隔の分布などとして使われる。

確率密度関数 : $f(x) = \lambda e^{-\lambda x} \quad (0 < x)$

分布関数 : $F(x) = 1 - e^{-\lambda x} \quad (0 < x)$

平均 : $\mu = \dfrac{1}{\lambda}$　　分散 ： $\sigma^2 = \dfrac{1}{\lambda^2}$



また、指数分布は無記憶性 $P(X > t + s | X > t) = P(X > s)$ を持っている。

例　$\lambda = 1$ の指数分布に従う確率変数 X が区間 $(2, 3]$ に入る確率

$$P(2 < X \leqq 3) = \int_2^3 e^{-x}\, dx = \left[-e^{-x} \right]_2^3 = e^{-2} - e^{-3} = 0.08554 \cdots = 0.0855$$

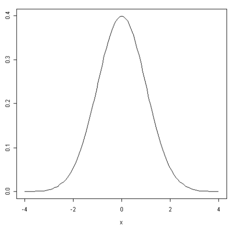
・正規分布 (パラメータは μ, σ^2)

統計で最も用いられる基本的な分布。身長・体重や誤差の分布などとして使われる。

$$\text{確率密度関数}: f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (-\infty < x < \infty)$$

$$\text{分布関数}: F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

平均 : μ　　分散 : σ^2



また、正規分布は $N(\mu, \sigma^2)$ と書くことが多い。特に $N(0,1)$ を標準正規分布と呼ぶ。

－正規分布の確率計算－

一般に正規分布の確率密度関数は定積分することができないので、コンピュータや近似値の表を使って計算する。「R」を使った計算は来週行うので、ここでは表を使った計算法を示す。

例　確率変数 X が正規分布 $N(5, 2^2)$ に従うとき、区間 $[2, 6]$ に入る確率

1．表は標準正規分布の確率の値のみなので、$z = \dfrac{x-\mu}{\sigma}$ で変換する。

$z = \dfrac{x-5}{2}$ で変換すると区間は $\left[\frac{2-5}{2}, \frac{6-5}{2}\right] = [-1.5, 0.5]$ となる。

つまり、正規分布 $N(5, 2^2)$ に従うとき、区間 $[2, 6]$ に入る確率と標準正規分布 $N(0, 1^2)$ に従うとき、区間 $[-1.5, 0.5]$ に入る確率が等しい。

2．表は $P(0 < z \leqq a)$ なので、0 からの確率になるように変形する。

$[-1.5, 0.5]$ を 0 で分割して $[-1.5, 0]$ と $[0, 0.5]$ の 2 つの確率の和と考える。また、正規分布は左右対称なので、$[-1.5, 0]$ と $[0, 1.5]$ の確率は等しい。

3．表の値を読んで確率の計算をする。

$$\begin{aligned} P(2 \leqq x \leqq 6) &= P(-1.5 \leqq x \leqq 0.5) \\ &= P(-1.5 \leqq x \leqq 0) + P(0 < x \leqq 0.5) \\ &= P(0 \leqq x \leqq 1.5) + P(0 < x \leqq 0.5) \\ &= 0.43319 + 0.19146 = 0.62465 = 0.625 \end{aligned}$$

数理統計学 R のインストールチェックと簡単な使い方

・R のインストールチェック

・R が起動するかチェック

・R を起動したとき、日本語が文字化けしていないかチェック

（日本語を使うときはフォントを MS Gothic にするのがお勧め。
編集 → GUI プリファレンスのフォントを変更）

・簡単な式（例．1+1）を入れて答の確認

・「demo(graphics)」と入力して、デモの動作確認

上記のことが全て動作すれば問題なし

・R の簡単な使い方

・$[0,1]$ の一様乱数を１０個発生

```
runif(10)
```

・データを入力してヒストグラムを書く

```
x <- c(11,13,17,20,25,27,28,30,31,33,32,27)
hist(x)
```
← 括弧の中身は何でも可

・一様乱数のヒストグラムを書く

```
x <- runif(100)
hist(x)
```
← 括弧の中身は乱数の数

・標準正規分布に従う乱数のヒストグラムを書く

```
x <- rnorm(100)
hist(x)
```
← 括弧の中身は乱数の数

・三角関数 $(\sin x)$ のグラフを描く

```
curve(sin(x),-5,5)
```

・標準正規分布のグラフを描く

```
curve(dnorm(x),-4,4)
```

・標準正規分布従う乱数のヒストグラムと標準正規分布のグラフを描く

```
x <- rnorm(100)
hist(x,xlim=c(-4,4),ylim=c(0,0.5),freq=F)
par(new="T")
curve(dnorm(x),-4,4,ylim=c(0,0.5))
```
← グラフの上書き命令（par(new="T") の行）

・おまけ　ヒストグラムの幅を固定

```
hist(x,breaks = seq(-5,5,0.5))
```
← -5 から 5 まで 0.5 刻みで

標準正規分布の確率 $P(0 < z \leqq a) = p$　　$a \to p$

a	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.00000	.00399	.00798	.01197	.01595	.01994	.02392	.02790	.03188	.03586
0.1	.03983	.04380	.04776	.05172	.05567	.05962	.06356	.06749	.07142	.07535
0.2	.07926	.08317	.08706	.09095	.09483	.09871	.10257	.10642	.11026	.11409
0.3	.11791	.12172	.12552	.12930	.13307	.13683	.14058	.14431	.14803	.15173
0.4	.15542	.15910	.16276	.16640	.17003	.17364	.17724	.18082	.18439	.18793
0.5	.19146	.19497	.19847	.20194	.20540	.20884	.21226	.21566	.21904	.22240
0.6	.22575	.22907	.23237	.23565	.23891	.24215	.24537	.24857	.25175	.25490
0.7	.25804	.26115	.26424	.26730	.27035	.27337	.27637	.27935	.28230	.28524
0.8	.28814	.29103	.29389	.29673	.29955	.30234	.30511	.30785	.31057	.31327
0.9	.31594	.31859	.32121	.32381	.32639	.32894	.33147	.33398	.33646	.33891
1.0	.34134	.34375	.34614	.34849	.35083	.35314	.35543	.35769	.35993	.36214
1.1	.36433	.36650	.36864	.37076	.37286	.37493	.37698	.37900	.38100	.38298
1.2	.38493	.38686	.38877	.39065	.39251	.39435	.39617	.39796	.39973	.40147
1.3	.40320	.40490	.40658	.40824	.40988	.41149	.41308	.41466	.41621	.41774
1.4	.41924	.42073	.42220	.42364	.42507	.42647	.42785	.42922	.43056	.43189
1.5	.43319	.43448	.43574	.43699	.43822	.43943	.44062	.44179	.44295	.44408
1.6	.44520	.44630	.44738	.44845	.44950	.45053	.45154	.45254	.45352	.45449
1.7	.45543	.45637	.45728	.45818	.45907	.45994	.46080	.46164	.46246	.46327
1.8	.46407	.46485	.46562	.46638	.46712	.46784	.46856	.46926	.46995	.47062
1.9	.47128	.47193	.47257	.47320	.47381	.47441	.47500	.47558	.47615	.47670
2.0	.47725	.47778	.47831	.47882	.47932	.47982	.48030	.48077	.48124	.48169
2.1	.48214	.48257	.48300	.48341	.48382	.48422	.48461	.48500	.48537	.48574
2.2	.48610	.48645	.48679	.48713	.48745	.48778	.48809	.48840	.48870	.48899
2.3	.48928	.48956	.48983	.49010	.49036	.49061	.49086	.49111	.49134	.49158
2.4	.49180	.49202	.49224	.49245	.49266	.49286	.49305	.49324	.49343	.49361
2.5	.49379	.49396	.49413	.49430	.49446	.49461	.49477	.49492	.49506	.49520
2.6	.49534	.49547	.49560	.49573	.49585	.49598	.49609	.49621	.49632	.49643
2.7	.49653	.49664	.49674	.49683	.49693	.49702	.49711	.49720	.49728	.49736
2.8	.49744	.49752	.49760	.49767	.49774	.49781	.49788	.49795	.49801	.49807
2.9	.49813	.49819	.49825	.49831	.49836	.49841	.49846	.49851	.49856	.49861

標準正規分布の確率 $P(0 < z \leqq a) = p$　　$p \to a$

p	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0.40	1.282	1.287	1.293	1.299	1.305	1.311	1.317	1.323	1.329	1.335
0.41	1.341	1.347	1.353	1.359	1.366	1.372	1.379	1.385	1.392	1.398
0.42	1.405	1.412	1.419	1.426	1.433	1.440	1.447	1.454	1.461	1.468
0.43	1.476	1.483	1.491	1.499	1.506	1.514	1.522	1.530	1.538	1.546
0.44	1.555	1.563	1.572	1.580	1.589	1.598	1.607	1.616	1.626	1.635
0.45	1.645	1.655	1.665	1.675	1.685	1.695	1.706	1.717	1.728	1.739
0.46	1.751	1.762	1.774	1.787	1.799	1.812	1.825	1.838	1.852	1.866
0.47	1.881	1.896	1.911	1.927	1.943	1.960	1.977	1.995	2.014	2.034
0.48	2.054	2.075	2.097	2.120	2.144	2.170	2.197	2.226	2.257	2.290
0.49	2.326	2.366	2.409	2.457	2.512	2.576	2.652	2.748	2.878	3.090

自由度 m の t 分布の両側 100α% 点

m	0.1	0.05	0.02	0.01
1	6.3137	12.706	31.821	63.656
2	2.9200	4.3027	6.9645	9.9250
3	2.3534	3.1824	4.5407	5.8408
4	2.1318	2.7765	3.7469	4.6041
5	2.0150	2.5706	3.3649	4.0321
6	1.9432	2.4469	3.1427	3.7074
7	1.8946	2.3646	2.9979	3.4995
8	1.8595	2.3060	2.8965	3.3554
9	1.8331	2.2622	2.8214	3.2498
10	1.8125	2.2281	2.7638	3.1693

2019年度 数理統計学 N科　小テスト解答用紙　2019.10. 9

・連続型確率の計算

次の確率の値を求めよ

1）区間 $[-2,4]$ の一様分布において、区間 $[1.2,3.3]$ に入る確率。

2）パラメータ $\lambda=\frac{1}{3}$ の指数分布において、区間 $[3,12]$ に入る確率。

3）$\mu=0,\sigma^2=1$ の標準正規分布 $N(0,1)$ において、区間 $[-1.89,-0.15]$ に入る確率。

2019年度神奈川工科大学 数理統計学 演習問題	学科	学年	組	学　籍　番　号	氏　　　名	

提出先：K3-3309号室前　19番のボックス　提出期限：１０月１０日（木）１７時頃まで