2018/01/31

藪友良

第 6 章の答え

1 $E[X]=1p+0(1-p)=p$、$E[X^2]=1^2p+0^2(1-p)=p$ となりますから、分散は $V(X)=p-p^2=p(1-p)$となります。

2 1)X は品質が良ければ 1 となるベルヌーイ確率変数とします（p=0.7）中古車の価値は 80X+30(1-X)ですから、その期待値は E[80X+30(1-X)]=80×0.7+30×0.3=65 万円（市場価格）。2)売り手は品質の良い車に 80 万円の価値があることを知っており、価格 65 万円では悪い品質の車だけを販売します。よって、中古車市場では悪い品質の車だけが取引され、良い品質の車は市場から無くなってしまいます（経済学では、このような市場をレモン市場といいます）。この問題を解決するためには、情報の非対称性を取り除く、つまり、売り手が買い手に中古車の正確な情報を開示する必要があります。

3 1) X～B(10,1/3)なら E[X]=10/3 です。E[10X-5(10-X)]=10E[X]-5(10-E[X])=10×10/3-5(10-10/3)=0 から適当に答えても平均 0 です。2) 10 問中 x 問正解する確率は、

$$P\{X = x\} = \frac{10!}{x!(10-x)!}\left(\frac{1}{3}\right)^{x}\left(\frac{2}{3}\right)^{10-x}$$

です。総得点が正とは（10x-5(10-x)>0）、x>50/15 のときですから、4 問以上に正しく回答すれば総得点は正です。

$$P\{X \geq 4\} = 1 - \sum_{i=0}^{3} P\{X = i\}$$

これを計算すると、P{X≥4}=0.44 です。

4 二項確率変数は相互に独立なベルヌーイ確率変数（1 となる確率は p、0 となる確率は 1-p）の和となります。しかし、相撲における試合結果は互いに独立ではないでしょう。調子の良し悪しがあり、前の試合の結果が今日の結果に影響するでしょう。また、相手の強さに応じて、勝率も一定ではなく、毎回変わると考えられます。以上から、勝利数 X を相互に独立なベルヌーイ確率変数の和と考えるのは厳密には正しくありません。

5　今日雨が降ると、明日も雨が降りやすくなるものです。また、時期によって降水確率も一定でないでしょう。こう考えると、雨の降る回数 X を相互に独立なベルヌーイ確率変数の和と考えるのは正しくありません[1]。

6　二項分布は相互に独立なベルヌーイ確率変数の和であることを思い出してください。ここで $X=X_1+X_2+\ldots+X_n$、$Y= Y_1+Y_2+\ldots+Y_m$ であり、X_i と Y_j は相互に独立なベルヌーイ確率変数（ともに p の確率で 1 となり、1-p の確率で 0 となる）とします。このとき、$X+Y= X_1+X_2+\ldots+X_n+Y_1+Y_2+\ldots+Y_m$ であり、各要素は相互に独立な n+m 個のベルヌーイ確率変数の和ですから、これは 2 項分布 B(n+m,p)に従います[2]。

7　1)
$$P\{X>10\}=P\left\{\frac{X-12}{2}>\frac{10-12}{2}\right\}=P\{Z>-1\}$$

また、分布表から $P\{Z>-1\}=P\{Z<1\}=0.8413$ となります。

2)
$$P\{X<8\}=P\left\{\frac{X-12}{2}<\frac{8-12}{2}\right\}=P\{Z<-2\}$$

また、分布表から $P\{Z<-2\}=P\{Z>2\}=1-P\{Z<2\}=1-0.9772=0.0228$ となります。

3)左辺は 12 で引かれていますから、標準化は両辺を 2 で割ります。

$$P\{|X-12|<3\}=P\left\{\left|\frac{X-12}{2}\right|<\frac{3}{2}\right\}=P\{|Z|<1.5\}$$

$P\{|Z|<1.5\}=P\{Z<1.5\}-P\{Z<-1.5\}=P\{Z<1.5\}-(1-P\{Z<1.5\})=0.9332-(1-0.9332)=0.8664$ となります。

8
$$P\{220<X<280\}=P\left\{\frac{220-230}{20}<\frac{X-230}{20}<\frac{280-230}{20}\right\}$$

$$=P\{-0.5<Z<2.5\}= P\{Z<2.5\}- P\{Z<-0.5\}$$

$$= P\{Z<2.5\}-(1-P\{Z<0.5\})=0.9938-(1-0.6915)=0.6853$$

9　X～B(14,1/2)ですから、1) X=4 の確率は

$$P\{X=4\}=\frac{14!}{4!(14-4)!}(1/2)^4(1/2)^{14-4}=0.0611$$

[1] 気候が急激に変化しているとき、過去のデータから雨の確率 p を推定するのは問題です。この場合、至近のデータだけを用いて p を推定する必要があります。

[2]厳密には、分布関数が一致することで証明します。

2) np=14(1/2)=7、npq=14(1/2)^2=3.5 から、二項分布は X～N（7,3.5）で正規近似できます。したがって、

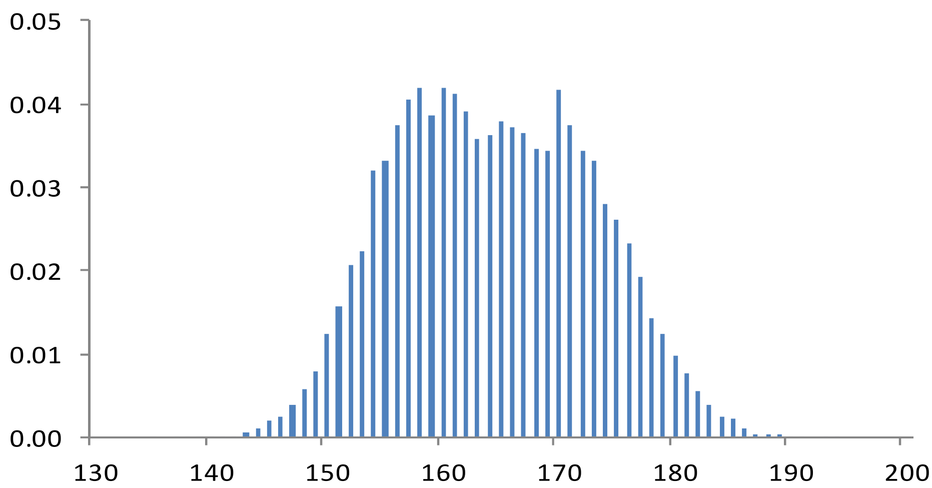
$$P\{3.5 < X < 4.5\} = P\left\{\frac{3.5-7}{\sqrt{3.5}} < \frac{X-7}{\sqrt{3.5}} < \frac{4.5-7}{\sqrt{3.5}}\right\}$$

$$= P\{-1.87 < Z < -1.34\} = 0.0594$$

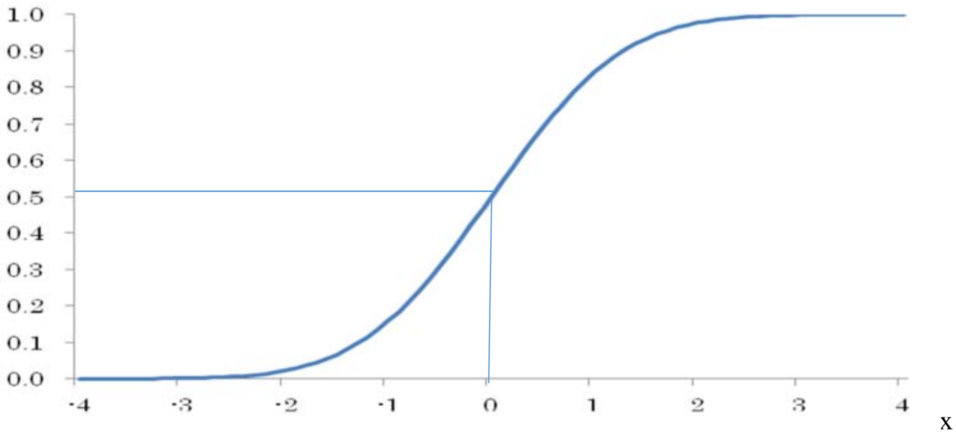
10 適切に設計された試験では、点数は正規分布に従うものです。もし 80 点で人数が極端に多ければ、何らかの不正が行われた可能性があります。この場合、80 点前後の答案を調べて、何らかの不正が行われた可能性を分析する必要があります。

11 男女全てを 1 つのグループとして身長の分布を図示しました（男女の身長の和ではありません）。男女全体の身長分布は、男性と女性の身長分布を混合させた分布（混合分布）となっています。頂点が 2 つあるのは、男性平均 170cm と女性平均 158cm で頂点が 2 つ形成されているためです。

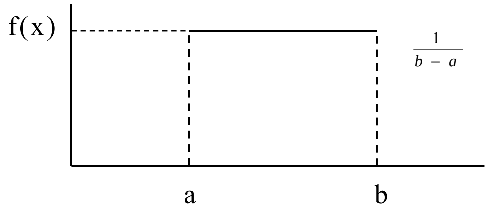
正規分布の性質から、正規確率変数の線形結合は正規分布に従いました。たとえば、男子の身長＋女子の身長は正規分布に従います。しかし、男子と女子を 1 つのグループとするということは、正規確率変数の線形結合ではないため、男女全員の身長分布は正規分布に従いません。



12 標準正規確率変数の累積分布関数は以下の通りです。x=0 で、確率は 50%となっています。これは標準正規分布が 0 を中心とした左右対称の分布だからです。



13 1)一様分布とは、全ての結果が同じ確率で生じる分布をいいます。ここで確率変数 X は連続確率変数であり、a から b の間の値を一定の確率でとりうるとしています。下図は、密度関数を示したものです。



2)密度関数は高さが $1/(b\text{-}a)$、幅が $b\text{-}a$ となります。面積=高さ×幅ですから、面積が 1 となることを確認できます。

3) a から b の区間の中点$(a+b)/2$ が期待値となります。直観的には、期待値は分布の中心ですから、期待値が中点であることは理解しやすいでしょう。

厳密には、積分を用いて期待値を求めます。確率変数が離散的な値をとるのであれば、期待値は取りうる値を確率で加重平均することで求めることができました。連続的確率変数の場合も、期待値は、取りうる値を確率で加重平均することで求めることができます。ただし、その確率は密度関数 $f(x)$であり、和 Σ は積分 ∫ を用いることになります。数式で書くと、確率変数 X が[a,b]区間の値をとるとき、期待値は

$$E[X]=\int_a^b xf(x)dx$$

となります。

一様確率変数の場合、確率変数 X は、$[a,b]$区間の範囲では $f(x)=1/(b-a)$としましたから、その期待値は

$$E[X] = \int_a^b x \frac{1}{(b-a)} dx = \frac{1}{(b-a)} \int_a^b x dx$$

となります(1/(b-a)は定数ですから、∫記号の外に出すことができます)。ここで上式に、

$$\int_a^b x dx = \left[\frac{x^2}{2}\right]_a^b = \left(\frac{b^2}{2} - \frac{a^2}{2}\right) = \frac{(a+b)(a-b)}{2}$$

を代入すれば、E[X]=(a+b)/2 を示せます。