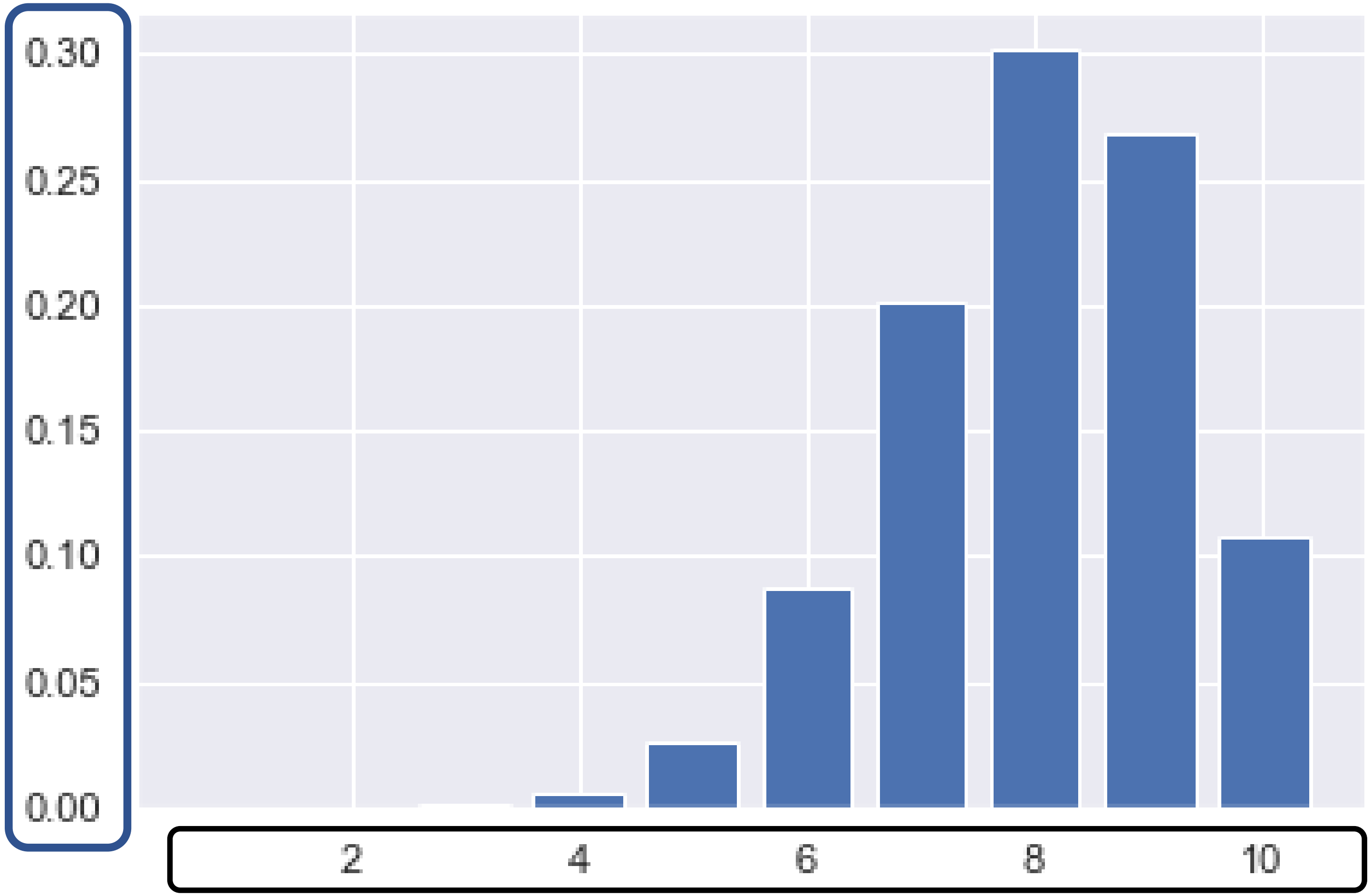
統計：二項分布とポアソン分布

確率分布とは

縦軸：Xが起こる確率

離散型：PMF$(X = x_i)$

連続型：PDF$(X = x_i)$



横軸：確率変数X

$(X = x_1, x_2, \cdots, x_n)$

二項分布

例1「平らなコインを10回投げて8回表が出る確率は？」

1回の試行で表が出る確率：$\frac{1}{2}$

求める確率 $= \binom{10}{8} (\frac{1}{2})^8 (1 - \frac{1}{2})^2$

例2「PKを10回やって8回ゴールが決まる確率は？」

1回のPKが決まる確率：p

求める確率 $= \binom{10}{8} p^8 (1 - p)^2$

二項分布

成功(確率p)または失敗(確率$1-p$)のどちらかになる試行をn回繰り返す
n回中，x回成功する確率$P(n,x)$を求める

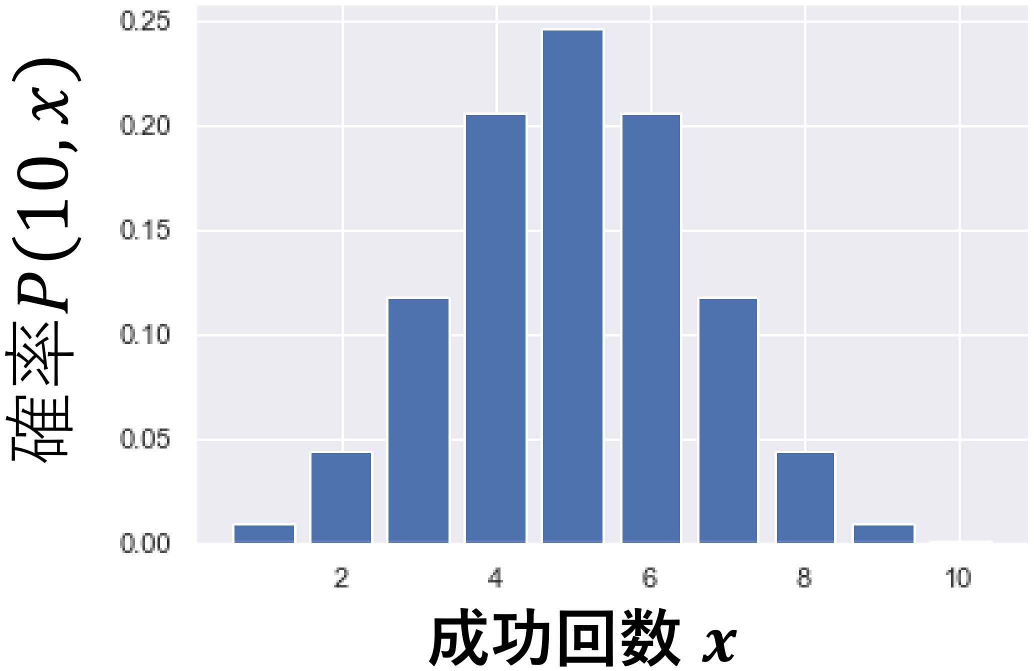
$$P(n,x)=\binom{n}{x}p^x(1-p)^{n-x}$$

n:試行回数　**x:成功する回数**

p: 1 回の試行で成功する確率

二項分布の確率質量関数(PMF)と呼ぶ

$n=10, p=\frac{1}{2}$ のときの確率分布



二項分布

二項分布の確率質量関数(PMF)

$$\mathrm{PMF}(n,x)=\binom{n}{x}p^x(1-p)^{n-x}$$

n:試行回数　x:成功する回数
p: 1回の試行で成功する確率

確率変数Xが二項分布に従うとき

・$E(X)=np$　・$V(X)=np(1-p)$

試行回数nがわかっている事象のカウントデータとかで使う！！

・コインをn回投げて，x回表がでる確率
・さいころをn回投げて，x回１がでる確率
・実験をn回試行して，x回成功する確率

ポアソン分布

二項分布：$\mathrm{PMF}(n, x) = \binom{n}{x} p^x (1-p)^{n-x}$

▶ **試行回数nがわかっていないとき**
（そもそも１回の試行を定義するのが難しい）

・1試合の中でゴールがx回決まる確率？
・1日に交通事故がx回起こる確率？
・1ヶ月に問い合わせの電話がx回来る確率？

でも，ある区間あたりに生起する数の期待値は分かっている

→ **ポアソン分布**

ポアソン分布

二項分布：$\mathrm{PMF}(n, x) = \binom{n}{x} p^x (1-p)^{n-x}$

$n \to \infty, p \to 0$　x回成功する数の期待値：$pn \to \lambda$

・成功確率が$p = \dfrac{\lambda}{n}$であるような試行を$n \to \infty$回繰り返す

ポアソンの少数の法則

$$\lim_{n \to \infty} \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \cdots = e^{-\lambda} \frac{\lambda^x}{x!}$$

⇦これがポアソン分のPMF

ポアソン分布

ポアソン分布の確率質量関数

$$\mathrm{PMF}(x) = e^{-\lambda}\frac{\lambda^x}{x!}$$

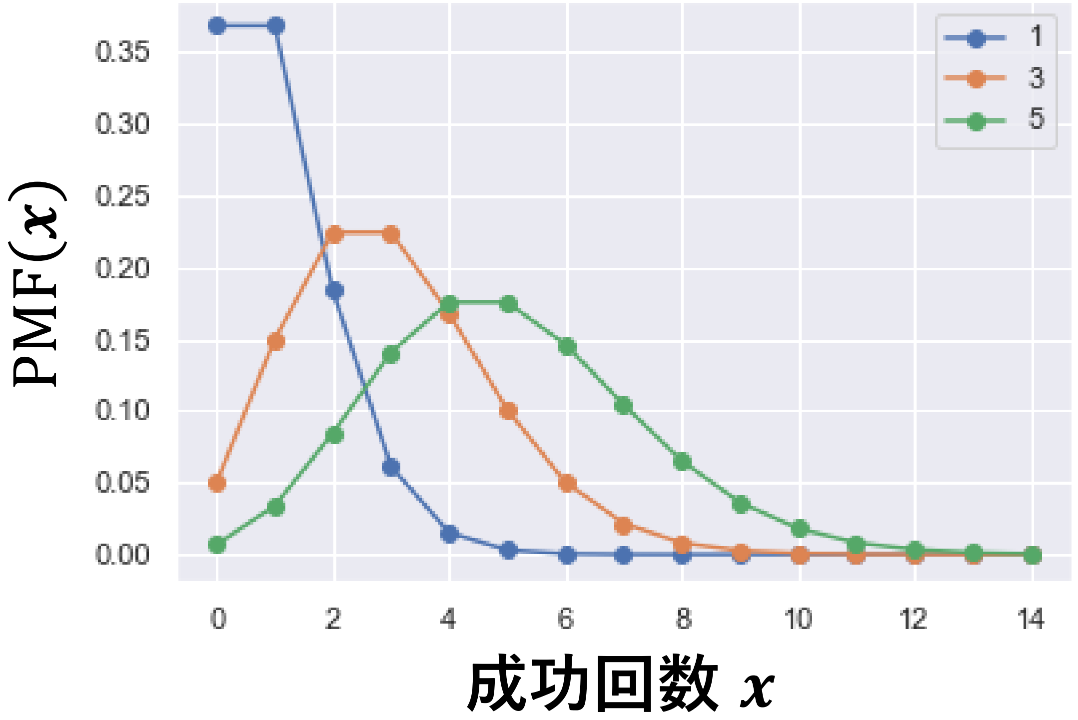
x:成功する回数

λ:成功する回数の期待値

確率変数Xがポアソン分布に従うとき

・$E(X) = \lambda$ ・$V(X) = \lambda$

$\lambda = 1,3,5$の時のポアソン分布

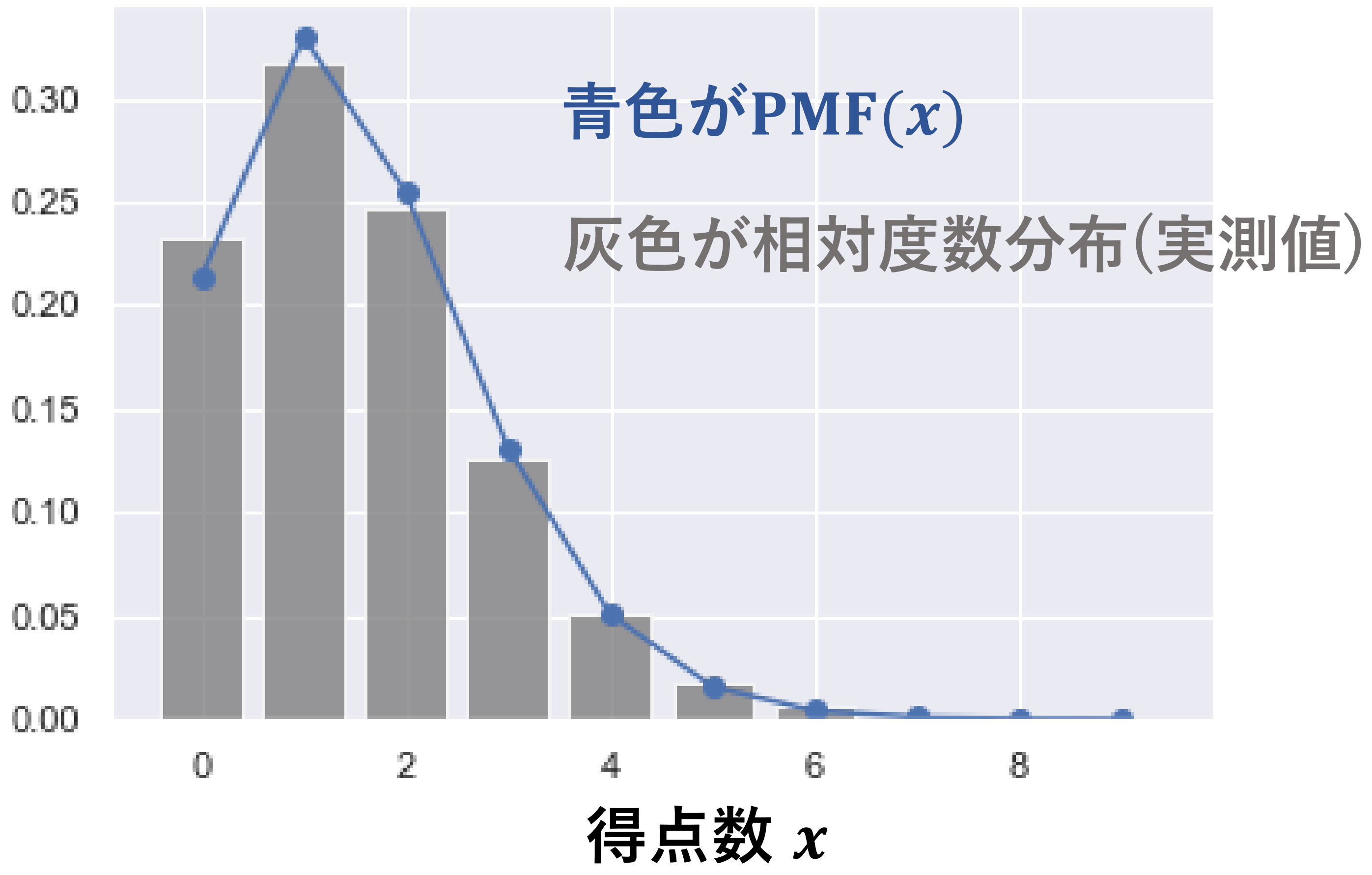


ポアソン分布

適用例「サッカー：1試合あたりのホームチームの得点数」

期待値$\lambda = 1.543$点

$$\mathrm{PMF}(x) = e^{-\lambda}\frac{\lambda^x}{x!}$$

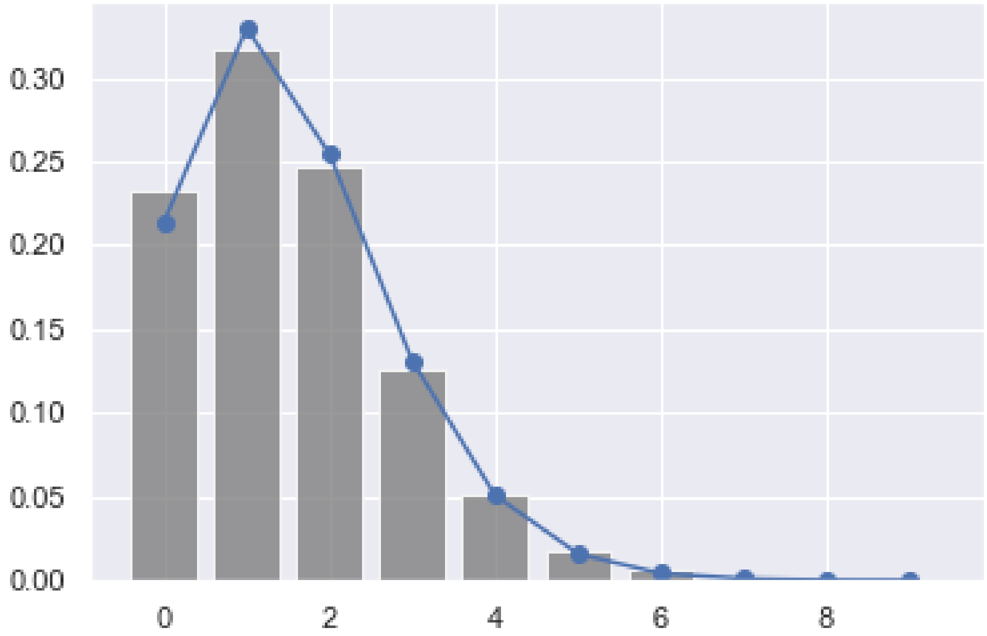


使用データ：https://www.kaggle.com/zaeemnalla/premier-league

ポアソン分布

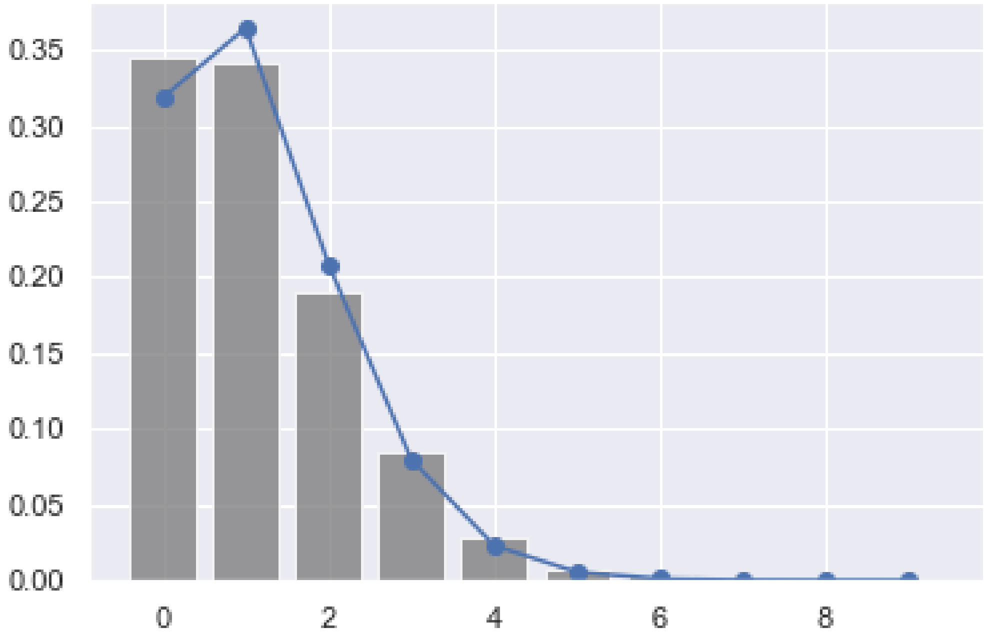
ホームとアウェイで比べてみる

ホームチームの得点数の分布



期待値$\lambda = 1.543$点

アウェイチームの得点数の分布



期待値$\lambda = 1.144$点