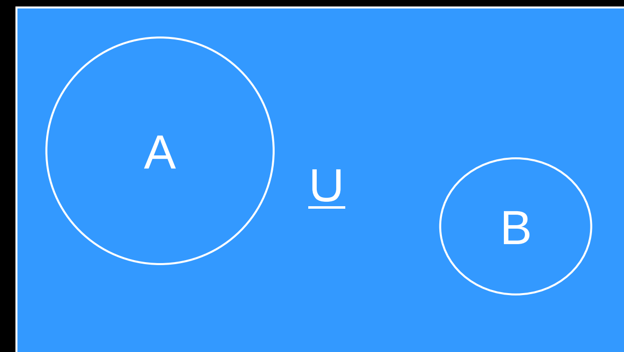
確率と統計

不規則なゆらぎ
平均的な振舞には規則がある

確率の定義（1）

- 事象（出来事）
  - 全集合（すべての出来事）：U
  - 空集合（何も起きない）：Φ
  - 背反事象（同時に起きない）：A･B＝Φ
- 相対頻度
  - 事象Aが起きる回数/全体の回数
  - 無限回の施行による極限→Aが起きる確率



$$P(A) = \lim \frac{m_{\mathrm{A}}}{n}$$

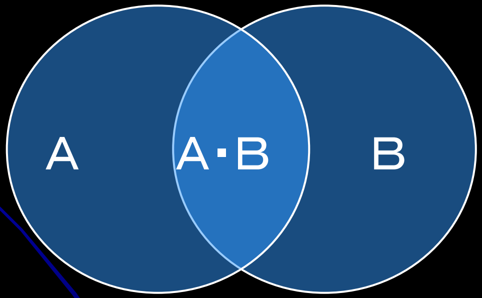
$$P(\Phi) = 0, \qquad P(U) = 1$$

$$A \bullet B = \Phi \Rightarrow P(A+B) = P(A) + P(B)$$

確率の定義（2）

- 結合確率
  - 排他的でない事象　Ａ　と　Ｂ
    - Ａ　と　Ｂ　が同時に起きる確率　P(A･B)
    - Ａ　と　Ｂ　の一方か両方が起きる確率　P(A＋B)

$$P(A + B) = P(A) + P(B) - P(A \cdot B)$$



例：　さいころを1回振るとき
Ａ＝偶数の目がで出る
Ｂ＝４以上の目が出る
Ａ･Ｂ＝４、６が出る
Ａ＋Ｂ＝２，４、５、６が出る

確率の定義(3)

- 条件付確率
  - 「Bの生起」が「Aの生起」に影響する
  - 「すでにBが起きた」とき「Aが起きる確率」:

$$P(A \mid B) = P(A \cdot B)/P(B)$$

- ベイズの定理

$$P(A \mid B) = P(B \mid A)P(A)/P(B)$$

- AとBが独立

$$P(A \mid B) = P(A)$$

$$P(A \cdot B) = P(A)P(B)$$

確率密度関数

- 確率変数
  - 不規則に起きる量を表す変数
  - 変数のとる値が離散的　vs　連続的

確率密度関数 p(x)

- 確率変数Xが x と x+dx の間の値として起きる確率 p(x) dx

$$\int_{-\infty}^{\infty} p(x)dx = 1$$

平均値、標準偏差

- 平均値

$$\bar{x} = E[X] = \frac{x_1 n_1 + x_2 n_2 + \ldots + x_j n_j}{n_1 + n_2 + \ldots + n_j} = x_1 P(x_1) + x_2 P(x_2) + \ldots + x_j P(x_j)$$

$$\bar{x} = E[X] = \sum_{k=1}^{j} x_k P(x_j) = \int_{-\infty}^{\infty} x\, p(x) dx$$

- 標準偏差　σ

$$\sigma^2 = E\left[(X - \bar{x})^2\right] = \frac{\sum_{k=1}^{j} n_j (x_j - \bar{x})^2}{\sum_{k=1}^{j} n_j} = \int_{-\infty}^{\infty} (x - \bar{x})^2 p(x) dx$$

２つの確率変数の和

- 平均値

$$E[X+Y]=\iint(x+y)p(x,y)dxdy=\iint xp(x,y)dxdy+\iint yp(x,y)dxdy=\bar{x}+\bar{y}$$

- 独立な量の標準偏差

$$E\left[\left((x+y)-(\bar{x}+\bar{y})\right)^2\right]=E\left[\left((x-\bar{x})+(y-\bar{y})\right)^2\right]$$
$$=E[(x-\bar{x})^2+(y-\bar{y})^2+2(x-\bar{x})(y-\bar{y})]$$
$$=E[(x-\bar{x})^2]+E[(y-\bar{y})^2]+E[2(x-\bar{x})(y-\bar{y})]$$
$$=\sigma_x^2+\sigma_y^2$$
$$\because E[2(x-\bar{x})(y-\bar{y})]=2\iint(x-\bar{x})(y-\bar{y})p(x)p(y)dxdy$$
$$=2\int(x-\bar{x})p(x)dx\times\int(y-\bar{y})p(y)dy=0$$

確率分布(1)

- 2項分布
- P(A)=p, P(B)=1-p

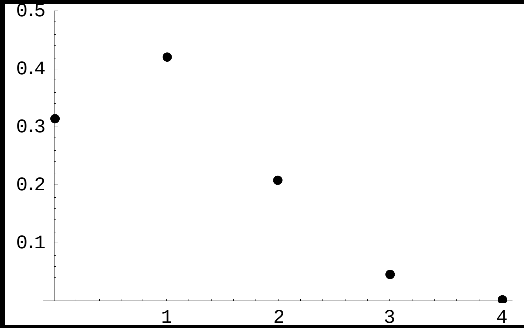
$$P_m = \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m}$$

平均値： $\overline{m} = \sum_{m=0,n} mP_m = np$

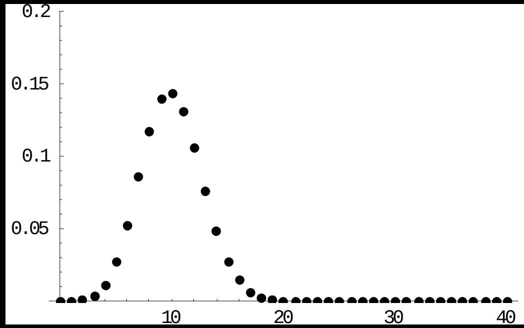
標準偏差： $\sigma = \sum (m-\overline{m})^2 P_m = \sqrt{np(1-p)}$

p=1/4

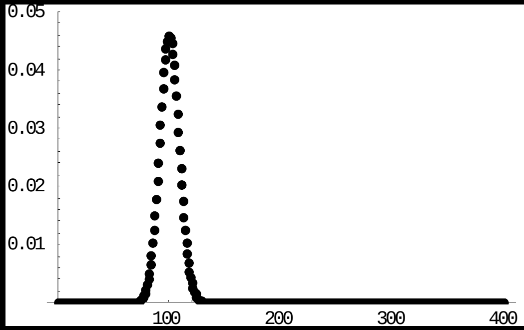
n= 4



n=40

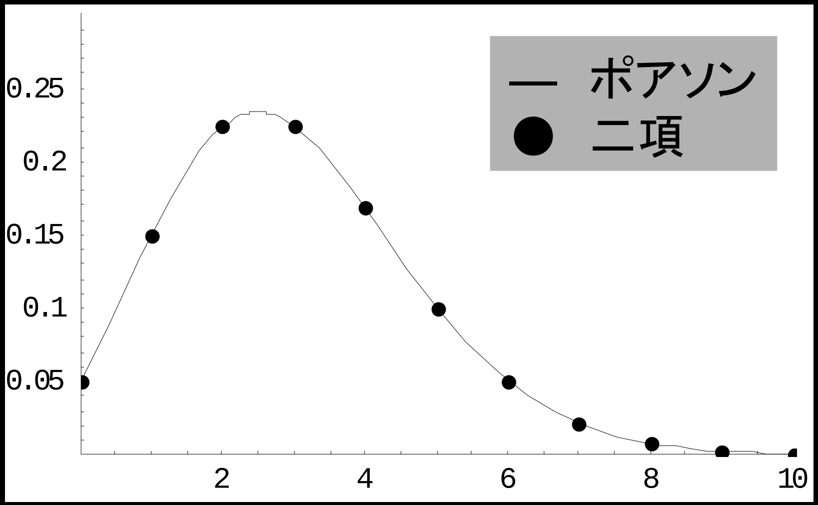


n=400



確率分布(2)

- ポアソン分布
  - 放射性崩壊、ショットノイズ…
  - n:非常に大きい、p:非常に小さい、平均:有限

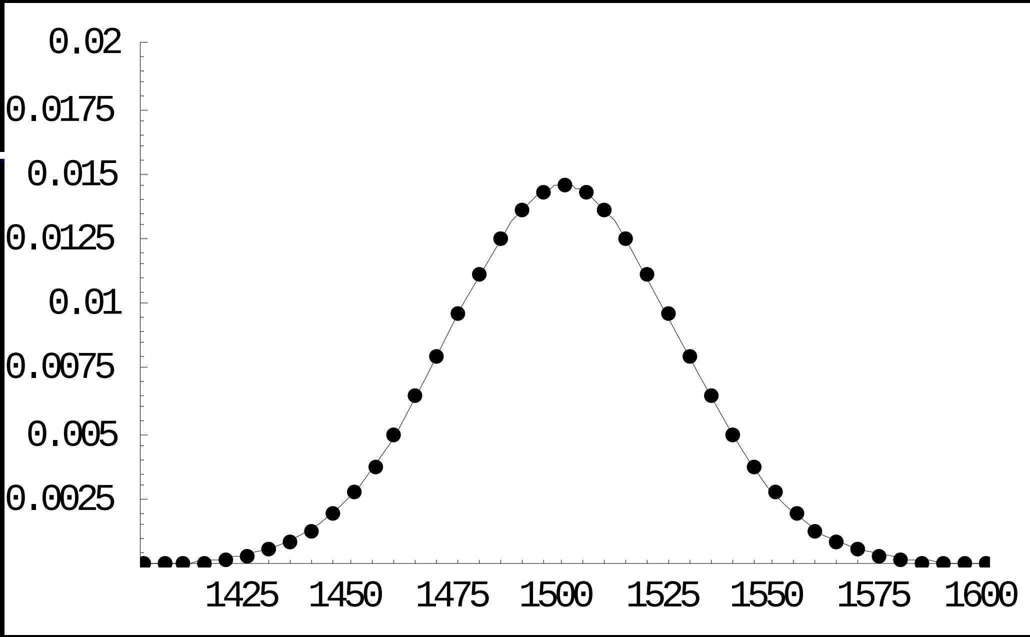


$$p_m = \frac{n(n-1)\ldots(n-m+1)}{m!} p^m (1-p)^{n-m} \cong \frac{n^m}{m!} p^m (1-p)^n$$

$$\cong \frac{(np)^m}{m!} \left(e^{-p}\right)^n = \frac{(np)^m}{m!} e^{-np} = \frac{\overline{m}^m}{m!} e^{-\overline{m}}$$

確率分布(3)

- ガウス(正規)分布
  - 測定誤差、白色ノイズ...
  - n, m 非常に大きい



$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\overline{m})^2}{2\sigma^2}}$$

ランダムプロセス

- 確率的な時系列
  - 時間の関数 y(t) が確率的に決まる
  - 確率密度 $p(y_1 t_1, y_2 t_2, \dots, y_n t_n)$
- 定常過程

$$p(y_1 t_1, y_2 t_2, \dots, y_n t_n)$$
$$=p(y_1(t_1+\tau), y_2(t_2+\tau), \dots, y_n(t_n+\tau))$$

  - 時間が経過しても確率密度関数が変化しない
  - 平均値などの統計量が時間的に変化しない

その他

- エルゴード性
  - 定常かつ時間平均と集団平均が同一