Дискретни разпределения

Разпределение на Бернули

$X \sim Ber(p)$, ако X има таблица на разпределение:

X	0	1
$\mathbb{P}(X = x_i)$	$q = 1 - p$	p

за някое $p \in [0, 1]$

Това може да тълкуваме като един опит с вероятност за успех p.

Тогава:

$\mathbb{E}[X] = 0 \cdot q + 1 \cdot p = p$
$\mathbb{D}[X] = p - p^2 = pq$
$g_X(s) = \mathbb{E}[s^X] = s^0 q + s^1 p = q + ps$

Биномно разпределение

$X \sim Bin(n, p)$, ако X е броят успехи измежду n независими бернулиеви опита с вероятност p за успех.

Тогава $X = \sum\limits_{j=1}^{n} X_j$, където $X_j \sim Ber(p)$

Твърдение: (свойства): Ако $X \sim Bin(n, p)$, то:

- $g_X(s) = (q + ps)^n$
- $\mathbb{E}[X] = np, \mathbb{D}[X] = npq$
- $\mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k}$ за $k = \overline{0, n}$
- **Доказателство:**
    - $X = \sum\limits_{j=1}^{n} X_j \overset{X_1 \overset{d}{=} X_j}{\Longrightarrow} g_X(s) = (g_{X_1})^n = (q + ps)^n$

- $\mathbb{E}[X] = \mathbb{E}[\sum_{j=1}^{n} X_j] = \sum_{j=1}^{n} \mathbb{E}[X_j] = np$, понеже $X_j \sim Ber(p)$
  $\mathbb{D}[X] = \mathbb{D}[\sum_{j=1}^{n} X_j] = \sum_{j=1}^{n} \mathbb{D}[X_j] = npq$, понеже $X_j \sim Ber(p)$
- Имаме, че $k! \cdot \mathbb{P}(X = k) = g_X^{(k)}(0)$
  $g_X^{(k)}(0) = n(n-1)\cdots(n-k+1)q^{n-k}p^k$
  Откъдето $\mathbb{P}(X = k) = \binom{n}{k}p^k q^{n-k}$

Геометрично разпределение

$X \sim Ge(p)$, ако X е броят неуспехи до първия успех в бернулиеви опити с вероятност p за успех, тоест

$$X = \min[n \geq 1 | \sum_{j=1}^{n} X_j = 1] - 1$$

Твърдение: (свойства): Ако $X \sim Ge(p)$, то:

- $\mathbb{P}(X = k) = pq^k$ за $k = \overline{0, n}$
- $g_X(s) = \dfrac{p}{1 - qs}$
- $\mathbb{E}[X] = \dfrac{q}{p}, \mathbb{D}[X] = \dfrac{q}{p^2}$
- **Доказателство:**
  - $\{X = k\} = \{X_1 = 0, \ldots, X_{k-1} = 0, X_k = 0, X_{k+1} = 1\}$, откъдето
    $\mathbb{P}(X = k) = \mathbb{P}(X_1 = 0, \ldots, X_k = 0, X_{k+1} = 1) \overset{iid}{=} q^k p$
  -

$$\begin{aligned} g_X(s) &= \mathbb{E}[s^X] \\ &= \sum_{k=0}^{\infty} s^k \mathbb{P}(X = k) \\ &= \sum_{k=0}^{\infty} s^k q^k p \\ &= p \sum_{k=0}^{\infty} (qs)^k \\ &\overset{|s|<1}{=} \frac{p}{1 - qs} \end{aligned}$$

- $\mathbb{E}[X] = g'_X(0) = \dfrac{q}{p}$

  $\mathbb{E}[X^2] = g''_X(0) + g'_X(0) = \dfrac{2q^2}{p^2} + \dfrac{q}{p}$

  $\mathbb{D}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \dfrac{2q^2}{p^2} + \dfrac{q}{p} - \dfrac{q^2}{p^2} = \dfrac{q}{p^2}$

Твърдение: (безпаметност): Ако $X \sim Ge(p)$, то $\mathbb{P}(X \geq n+m | X \geq n) = \mathbb{P}(X \geq m)$ за всяко $n, m \geq 0$

- **Доказателство:**
  - Понеже $\mathbb{P}(X \geq k) = \sum_{j=k}^{\infty} pq^j = pq^k \sum_{j=0}^{\infty} q^j = \dfrac{pq^k}{1-q} = q^k$

$$\begin{aligned} \mathbb{P}(X \geq n+m | X \geq n) &= \frac{\mathbb{P}(X \geq n+m, X \geq n)}{\mathbb{P}(X \geq n)} \\ &= \frac{\mathbb{P}(X \geq n+m)}{\mathbb{P}(X \geq n)} \\ &= \frac{q^{n+m}}{q^n} \\ &= q^m \\ &= \mathbb{P}(X \geq m) \end{aligned}$$

Отрицателно биномно разпределение

$X \sim NB(r, p)$, ако X е броят неуспехи до r-тия успех в бернулиеви опити с вероятност p за успех, тоест

$X = \min[n \geq 1 | \sum_{j=1}^{n} X_j = r] - r$

Твърдение: Ако $X \sim NB(r, p)$, то $X = \sum_{j=1}^{r} Y_j$, където $Y_j \sim Ge(p)$ са независими в съвкупност.

- **Доказателство:**

- $$\begin{aligned} X &= \min\left[n \geq 1 \middle| \sum_{j=1}^{n} X_j = r\right] - r \\ &= \min\left[n \geq 1 \middle| \sum_{j=1}^{n} Y_j = 1\right] + \cdots + \min\left[n \geq 1 \middle| \sum_{j=1}^{n} Y_j = 1\right] - r \\ &= \sum_{j=1}^{r} Y_j \end{aligned}$$
- Алтернативно, можем да докажем, че $g_X(s) = (g_Y(s))^r$, където $Y \sim Ge(p)$
$$g_X(s) = \mathbb{E}[s^X] = \mathbb{E}\left[s^{\sum_{j=1}^{r} Y_j}\right] = \mathbb{E}\left[\prod_{j=1}^{r} s^{Y_j}\right] \overset{Y_j \perp Y_i}{=} \prod_{j=1}^{r} \mathbb{E}[s^{Y_j}] = (g_Y(s))^r$$
- Алтернативно, можем да проверим, че $X = X_1 + X_2$, където $X_j \sim Ge(p)$ са независими в съвкупност.

Твърдение: (свойства): Ако $X \sim NB(r, p)$, то:

- $g_X(s) = \left(\dfrac{p}{1 - qs}\right)^r$
- $\mathbb{E}[X] = \dfrac{rq}{p}, \mathbb{D}[X] = \dfrac{rq}{p^2}$
- $\mathbb{P}(X = k) = \dbinom{k + r - 1}{k} p^r q^k$ за $k = \overline{0, n}$
- **Доказателство:**
  - $$g_X(s) = g_{X_1 + \cdots + X_r}(s) \overset{iid}{=} \prod_{j=1}^{r} g_{X_1}(s) = \left(\frac{p}{q - qs}\right)^r$$
  - $\mathbb{E}[X] = g'_X(0) = \dfrac{rq}{p}$
  $\mathbb{E}[X^2] = g''_X(0) + g'_X(0) = \dfrac{2rq^2}{p^2} + \dfrac{rq}{p}$
  $\mathbb{D}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \dfrac{2rq^2}{p^2} + \dfrac{rq}{p} - \dfrac{r^2q^2}{p^2} = \dfrac{rq}{p^2}$
  Или от свойствата на очакването и дисперсията.
  - $k! \cdot \mathbb{P}(X = k) = g_X^{(k)}(0)$, откъдето следва директно

Поасоново разпределение

$X \sim Poi(\lambda)$, ако X е броят на събития в единичен интервал, където събитията са независими и имат средно $\lambda > 0$ събития на интервал и $\mathbb{P}(X = k) = \dfrac{\lambda^k}{k!} e^{-\lambda}$

Проверка за коректност:

$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

Твърдение: (свойства): Ако $X \sim Poi(\lambda)$, то:

- $g_X(s) = e^{\lambda(s-1)}$
- $\mathbb{E}[X] = \lambda, \mathbb{D}[X] = \lambda$
- **Доказателство:**
    - $g_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} s^k \dfrac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \dfrac{(\lambda s)^k}{k!} = e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)}$
    - $\mathbb{E}[X] = g'_X(1) = \lambda$
      $\mathbb{E}[X^2] = g''_X(1) + g'_X(1) = \lambda^2 + \lambda$
      $\mathbb{D}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

Твърдение: Ако $X_j \sim Poi(\lambda)$, то $Y = \sum_{j=1}^{n} X_j \sim Poi(n\lambda)$

Теорема: (на Поасон): Нека $X_n \sim Bin(n, p_n)$ и $p_n = \dfrac{\lambda}{n} + \dfrac{V_n}{n}$, където $\lim_{n\to\infty} |V_n| = 0$.

Тогава

$$\mathbb{P}(X_n = k) \xrightarrow{n\to\infty} \frac{\lambda^k}{k!} e^{-\lambda}$$

- Тоест $X_n \xrightarrow[n\to\infty]{d} X \sim Poi(\lambda)$
- **Доказателство:**
    - $\lim\limits_{n\to\infty} \mathbb{P}(X_n) = \frac{\lambda^k}{k!} e^{-\lambda}$ е вярно, ако $\lim\limits_{n\to\infty} g_{X_n}(s) = g_X(s)$ за $X \sim Poi(\lambda)$
    - $g_{X_n}(s) = (q_n + p_n s)^n =$
      $\left(1 - \dfrac{\lambda}{n} - \dfrac{V_n}{n} + (\dfrac{\lambda}{n} + \dfrac{V_n}{n}) s\right)^n =$
      $\left(1 + \dfrac{\lambda + V_n}{n}(s-1)\right)^n \xrightarrow{n\to\infty} e^{\lambda(s-1)} = g_X(s)$

Хипергеометрично разпределение

$X \sim HG(N, M, n)$, ако X е броят на изтеглените маркирани обекти от общо изтеглени n обекта от всичките N обекта, където M от тях са маркирани.

Тогава $X = \sum\limits_{j=1}^{n} X_j$, където $X_j \sim Ber(\frac{M}{N})$

Твърдение: (свойства): Ако $X \sim HG(N, M, n)$, то:

- $\mathbb{P}(X = k) = \dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$ за $k = \overline{0, n}$
- $\mathbb{E}[X] = \dfrac{nM}{N}, \mathbb{D}[X] = \dfrac{nM(N-M)(N-n)}{N^2(N-1)}$
- **Доказателство:**
  - $\mathbb{P}(X = k) = \dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$, понеже имаме $\binom{N}{n}$ възможности за изтегляне на n обекта от N обекта, $\binom{M}{k}$ възможности за изтегляне на k маркирани обекта от M маркирани обекта и $\binom{N-M}{n-k}$ възможности за изтегляне на $n-k$ немаркирани обекта от $N-M$ не маркирани обекта.
  - $\mathbb{E}[X] = \mathbb{E}[\sum\limits_{j=1}^{n} X_j] = \sum\limits_{j=1}^{n} \mathbb{E}[X_j] = \sum\limits_{j=1}^{n} \dfrac{M}{N} = \dfrac{nM}{N}$
    $\mathbb{D}[X] = \mathbb{D}[\sum\limits_{j=1}^{n} X_j] = \sum\limits_{j=1}^{n} \mathbb{D}[X_j] = \sum\limits_{j=1}^{n} \dfrac{M(N-M)}{N^2} \dfrac{N-n}{N-1} = \dfrac{nM}{N} \dfrac{N-M}{N} \dfrac{N-n}{N-1}$