Seminar W2 – 831

→ Bernoulli trial : sequence of independent "coin flips"

→ there are two possibilities: success (A), failure ($\bar{A}$)

→ the trials are independent

→ success has the same probability (denoted by p in every trial)

→ We will discuss two models:

- **Binomial model** : n Bernoulli trials

  → X = the number of successes in these n trials

$$X \sim B(n,p) \qquad X \begin{pmatrix} k \\ C_n^k \cdot p^k \cdot (1-p)^{n-k} \end{pmatrix}_{k=\overline{0,n}}$$

$$\forall k \in \{0,1,2,\dots,n\} : \; P(X=k) = C_n^k \cdot p^k \cdot (1-p)^{n-k}$$

- **Negative binomial model** : sequence of trials (a countably infinite number of trials)

  $\downarrow$ $n=1$

  Geometric model

$X =$ "number of failures before the n^{th} success"

$$X \sim NB(n, p)$$

$$X \begin{pmatrix} k \\ C^{k}_{n+k-1} \cdot p^{n} \cdot \underbrace{(1-p)}_{=:q}{}^{k} \end{pmatrix}$$

Exercise 1. A computer virus has entered a system with a very large number of files. A computer manager identifies the type of the virus, thereby learning that each file is independently damaged with probability 0.2. Next, the manager runs a program to check the condition of each file. Find the probability that

(a) 2 of the first 10 scanned files are damaged;

(b) at least 3 of the first 20 scanned files are damaged;

(c) at least 19 files are scanned before 10 undamaged files are found.

(a) 10 trials Binomial model

$n = 10$ success = "one file is damaged"

$X =$ "number of damaged files within the first 10"

$$P(X=2) = C^{2}_{10} \cdot (0.2)^2 \cdot (0.8)^8 \approx 0.301$$

(b) $n = 20$ Binomial model

$X =$ "number of damaged files within the first 20"

$$P(X \geq 3) = 1 - P(X < 3) = 1 - \left(P(X=0) + P(X=1) + P(X=2)\right)$$

$$= 1 - \left(C_{20}^{0} \cdot (0.2)^0 \cdot (0.8)^{20} + C_{20}^{1} \cdot (0.2)^1 \cdot (0.8)^{19} + C_{20}^{2} \cdot (0.2)^2 \cdot (0.8)^{18} \right)$$

$$= 1 - \left((0.8)^{20} + 20 \cdot 0.2 \cdot (0.8)^{19} + 110 \cdot (0.2)^2 \cdot (0.8)^{18} \right) =$$

$$\simeq 1 - 0.148 = 0.852$$

(c) P("at least 19 are scanned before 10 undamaged are found")

X = "number of failures before the n^{th} success"

Y = "no of scanned before 10 undamaged are found"

$$P(Y \geq 19)$$

Z = "no of damaged files before 10 undamaged are found"

$$P(Y \geq 19) = P(Z \geq \alpha), \quad \alpha = ?$$

What does it mean that $Z = 10$?

$$\underbrace{\qquad\qquad\qquad\qquad}_{\substack{10 \text{ damaged} \\ 9 \text{ undamaged}}} \ \uparrow \ 10^{th} \text{ undamaged}$$

$$\Rightarrow Y = 19 \iff Z = 10$$

$\Rightarrow \quad Y \geq 19 \iff Z \geq 10$

$P(Y \geq 19) = P(Z \geq 10)$

X = "number of failures before the n^{th} success"

Z = " no of damaged files before 10 undamaged are found"

$\Rightarrow \quad Z \sim NB(n, p)$

$\Rightarrow n = 10 \Rightarrow p = 0.8$

$\Rightarrow Z : \begin{pmatrix} k \\ C_{9+k}^{k} \cdot (0.8)^{10} \cdot (0.2)^{k} \end{pmatrix}$

$\Rightarrow \quad P(Z \geq 10) = 1 - P(Z < 10) =$

$= 1 - P(Z \leq 9) = 1 - \sum_{k=0}^{9} C_{9+k}^{k} (0.8)^{10} \cdot (0.2)^{k}$

$= 1 - cdf(9)$

Exercise 3. A darts player aims at the bullseye, which has a diameter of 1 cm. The distance from the center of the dartboard to the hit point of an arrow thrown by the player follows the $Beta(a, b)$ distribution, where $a, b > 0$, with a mean value of $\frac{3}{4}$ cm and a standard deviation of $\frac{\sqrt{15}}{20}$ cm. We assume that the throws are independent.

(a) Find the probability that the player hits the bullseye on a throw.

(b) Find the mean value of the number of throws before the player hits the bullseye the first time.

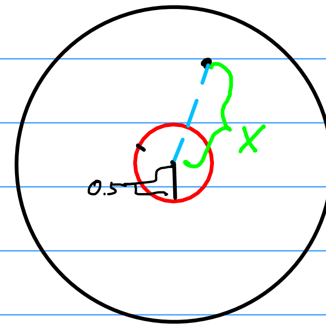
(c) Find the probability that the player hits the bullseye 2 times out of 10 throws.

Recap. *If $X \sim Beta(a, b)$, $a, b > 0$, then its pdf is:*

$$f_X(x) = \frac{1}{B(a,b)} \cdot x^{a-1}(1-x)^{b-1}$$

for $x \in [0,1]$ (and 0 otherwise). We have:

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx, \ a, b > 0$$



Sol.:

$$E(X) = \frac{3}{4} \qquad \sigma = \sqrt{V(X)} = \frac{\sqrt{15}}{20}$$

$$E(X) = \int_{\mathbb{R}} x \cdot f_X(x) \, dx =$$

$$= \int_0^1 x \cdot \frac{1}{B(a,b)} \cdot x^{a-1}(1-x)^{b-1} \, dx =$$

$$= \frac{1}{B(a,b)} \underbrace{\int_0^1 x^a \cdot (1-x)^{b-1} \, dx}_{B(a+1,b)} =$$

$$= \frac{B(a+1,b)}{B(a,b)} = \frac{a}{a+b} \qquad \Rightarrow E(X) = \frac{a}{a+b} = \frac{3}{4}$$

$$\left(5. \ B(a,b) = \frac{a-1}{a+b-1} B(a-1,b), \text{ for every } a > 1, b > 0; \right)$$

$$V(X) = E(X^2) - E(X)^2$$

using wikipedia

$$= \int_{\mathbb{R}} x^2 f_X(x) \, dx - \left(\frac{a}{a+b}\right)^2 = \frac{ab}{(a+b)^2(a+b+1)}$$

$$\Rightarrow \begin{cases} \dfrac{a}{a+b} = \dfrac{3}{4} \\ \dfrac{ab}{(a+b)^2 \cdot (a+b+1)} = \dfrac{15}{400} \end{cases}$$

$$\frac{a}{a+b} = \frac{3}{4} \Rightarrow 4a = 3a + 3b \Rightarrow a = 3b$$

$$\Rightarrow \frac{3b^2}{16b^2 \cdot (4b+1)} = \frac{15}{400}$$

$$\Rightarrow \overset{100}{\cancel{400}} \cdot \cancel{3} = \overset{5}{\cancel{15}} \cdot \overset{4}{\cancel{16}} (4b+1)$$

$$\Rightarrow 4b+1 = 5 \Rightarrow b = 1 \Rightarrow a = 3$$

$$\Rightarrow X \sim \text{Beta}(3,1)$$

a) P("the player hits the bullseye") =

$$= P\left(X \leq \frac{1}{2}\right) = F_X\left(\frac{1}{2}\right) = \int_0^{1/2} \frac{1}{B(3,1)} \cdot x^2 \cdot (1-x)^0 \, dx =$$

$$= \frac{1}{B(3,1)} \int_0^{1/2} x^2 \, dx = \frac{1}{B(3,1)} \cdot \left(\frac{1}{3} x^3\right)\Bigg|_0^{1/2} = \frac{1}{24 B(3,1)}$$

$$B(3,1) = \int_0^1 x^{3-1} \cdot (1-x)^{1-1} \, dx = \int_0^1 x^2 \, dx = \frac{1}{3}$$

$$\Rightarrow P\left(X \leq \frac{1}{2}\right) = \frac{1}{8}$$

(b) Y = "number of throws before the first success"

$$\Rightarrow Y \sim NB(1, p) = Geo(p) = Geo\left(\frac{1}{8}\right)$$

(c) $P(Z = 2)$

Z = "n° of hits out of 10 throws"

$$\Rightarrow Z \sim B\left(10, \frac{1}{8}\right)$$