Seminar W2 - 832

→ Bernoulli trials

- two possible outcomes: A (success), $\overline{A}$ (failure)
- independent
- same probability as success

→ **Binomial model**: n Bernoulli trials

X = "number of successful trials out of n"

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

→ **Negative binomial model**: a countably infinite number of trials

$n=1$ case → Geometrical model

X = "number of failures occurring before the nth success"

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

$$X \begin{pmatrix} k \\ C_{n+k-1}^k \cdot p^n \cdot \underbrace{(1-p)}_{"q"}{}^k \end{pmatrix}_{k \in \mathbb{N}}$$

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) A = "the file is damaged" $\Rightarrow P(A) = p = 0.2$

$\bar{A}$ = "the file is undamaged"

$n = 10$

X = "the number of damaged files in the first 10 scanned files"

$P(X=2) = ?$

We are using the binomial model, therefore

$X \sim B(10, 2)$

$\Rightarrow P(X=k) = C_n^k \cdot p^k \cdot q^{n-k}$

$P(X=2) = C_{10}^2 \cdot (0{,}2)^2 \cdot (0{,}8)^8 =$

$= \frac{10 \cdot 9}{1 \cdot 2} \cdot 0{,}04 \cdot \frac{2^{24}}{10^8}$

(b) $n = 20$, $X \sim B(20, 0.2)$

$P(X \geq 3) = ?$

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

$+ P(X=1) + P(X=2)\big)$

$P(X=0) = C_{20}^{0} \cdot (0.2)^0 \cdot (0.8)^{20} = (0.8)^{20} \approx 0.011$

$P(X=1) = C_{20}^{1} \cdot (0.2)^1 \cdot (0.8)^{19} \approx 0.057$

$P(X=2) = C_{20}^{2} \cdot (0.2)^2 \cdot (0.8)^{18} \approx 0.136$.

(c) X = " no of failures before the nth success "
$\approx NB(n,p)$

What we need: number of trials before the 10th failure"
(it needs to be at least 19)

number of trials = number of successes + no of failures

Y = number of damaged before the 20th undamaged is found

~~$P(Y \geq 9)$~~

if 8 damaged before 10th undamaged:

$\underbrace{\qquad\qquad}_{8\,d,\ 9\,u}$ ↑ 10th un

if 9 damaged before 10th undamaged

$\Rightarrow$ we need at least 10 damaged before 10th undamaged

$\Rightarrow P(Y \geq 10)$

Y = number of damaged before the 10th undamaged is found

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

Here success = undamaged

$n = 10 \quad p = 0.8$

$\Rightarrow Y \sim NB(10, 0.8)$

$$Y \begin{pmatrix} k \\ C_{n+k-1}^{k} \cdot p^{n} \cdot q^{k} \end{pmatrix} \Rightarrow Y \begin{pmatrix} k \\ C_{9+k}^{k} \cdot 0.8^{10} \cdot 0.2^{k} \end{pmatrix}$$

$$P(Y \geq 10) = 1 - P(Y < 10) =$$

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

$$\shortparallel$$

$$\frac{(9+k)!}{k! \cdot 9!}$$

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

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

for $x > 0$ (and 0 otherwise). We have:

$$\Gamma(a) = \int_0^{\infty} x^{a-1} e^{-x} dx, \; a > 0$$

Exercise 2. A user spends X minutes on a certain website, where

$$X \sim Gamma(a, b), \; a, b > 0$$

with a mean value of 12 minutes and a standard deviation of 6 minutes. Find the probability that the user spends at most 10 minutes on the website, given that he spends at least 1 minute.

Sol.: $E(X) = 12 \qquad \sigma = 6 \Rightarrow V(X) = 36$

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

$$= \int_{-\infty}^{0} x f_X(x) \, dx + \int_0^{\infty} x f_X(x) \, dx =$$

$$= \int_0^{\infty} x f_X(x) \, dx = \int_0^{\infty} x \cdot \frac{1}{\Gamma(a) b^a} \cdot x^{a-1} \cdot e^{-\frac{x}{b}} dx =$$

$$= \frac{1}{\Gamma(a) b^a} \int_0^{\infty} x^a \cdot e^{-\frac{x}{b}} dx \overset{\frac{x}{b} = y}{=} \frac{1}{\Gamma(a) b^a} \cdot \int_0^{\infty} y^a \cdot b^a \cdot e^{-y} \cdot b \, dy$$

$$= \frac{b}{\Gamma(a)} \underbrace{\int_0^{\infty} y^a e^{-y} dy}_{\Gamma(a+1)} = \frac{b \, \Gamma(a+1)}{\Gamma(a)} = ab$$

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

$$E(X^2) = \int_{\mathbb{R}} x^2 f_X(x)\, dx =$$

$$= \frac{1}{\Gamma(a) b^a} \int_0^\infty x^{a+1} e^{-\frac{x}{b}} \, dx \overset{\frac{x}{b}=y}{=}$$

$$= \frac{1}{\Gamma(a) b^a} \int_0^\infty b^{a+1} \cdot y^{a+1} e^{-y} \, b \, dy =$$

$$= \frac{b^2}{\Gamma(a)} \underbrace{\int_0^\infty y^{a+1} e^{-y} \, dy}_{\Gamma(a+2)} =$$

$$= \frac{b^2 \Gamma(a+2)}{\Gamma(a)} = \frac{b^2 (a+1) \Gamma(a+1)}{\Gamma(a)} = \frac{b^2 (a+1) \cdot a \Gamma(a)}{\Gamma(a)} =$$

$$= b^2 (a+1) \cdot a$$

$$V(X) = b^2 \cdot (a+1) a - a^2 b^2 = a b^2 (a+1-a) = a b^2 .$$

$$\Rightarrow \quad E(X) = ab \qquad V(X) = ab^2$$

We know : $E(X) = 12 \quad V(X) = 36$

$$\Rightarrow \begin{cases} ab = 12 \\ ab^2 = 36 \end{cases} \Rightarrow b = 3 \quad a = 4$$

$$\Rightarrow X \sim \text{Gamma}(4,3)$$

$\cancel{P(4 \le X \le 10) =}$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \qquad \text{(Bayes' law)}$$

$$P(X \leq 10 \mid X \geq 1) = \frac{P(1 \leq X \leq 10)}{P(X \geq 1)} = \frac{F_X(10) - F_X(1)}{1 - F_X(1)}$$

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\, dx$$

$$= \frac{\text{cdf Gamma}(4,3)\,(10) - \text{cdf Gamma}(4,3)\,(1)}{1 - \text{cdf Gamma}(4,3)\,(1)}$$