MATH 425 10/26/2022

The negative distributions

Any time we have a scenario with repeated Bernoulli trials (independent or not) we can count the number of attempts for the trial to succeed (or, more generally, to succeed r times).

Negative hypergeometric distribution: N balls, m of which are red. I keep drawing balls without replacement. How many balls do I have to draw to get r red balls? ($r \le m$).

Negative binomial distribution

① The case of $r=1$. Performing <u>independent</u> Bernoulli trials with probability p of success. N_p = number of trials it takes to succeed once. This is called the <u>geometric distribution</u>.

values : $\{1,2,3,4,\ldots\}\,(\cup\{\infty\})$

↖ probability 0, so it does not matter

<u>Probability mass function:</u>

$$P(N_p = 1) = p$$

$$P(N_p = 2) = (1-p)p = qp \quad \cdots \quad P(N_p = n) = (1-p)^{n-1}p = q^{n-1}p$$

Check:

$$\sum_{n=1,2,3,\dots} P(N_p = n) = p + p(1-p) + p(1-p)^2 + \cdots =$$

$$= p\Big(\underbrace{1 + (1-p) + (1-p)^2 + \cdots}_{\text{geometric series}}\Big) = p\,\frac{1}{1-(1-p)} = 1$$

Calculating the expectation $E(N_p)$:

$$E(N_p) = \sum_{n=1}^{\infty} n(1-p)^{n-1}p = p\begin{pmatrix} 1 + (1-p) + (1-p)^2 + \cdots \\ + (1-p) + (1-p)^2 + \cdots \\ + (1-p)^2 + \cdots \\ + \cdots \end{pmatrix}$$

$\frac{1}{p}$

$\frac{1-p}{p}$

$\frac{(1-p)^2}{p}$

$$= 1 + (1-p) + (1-p)^2 + \cdots = \frac{1}{1-(1-p)} = \frac{1}{p}$$

$$\boxed{E(N_p) = \frac{1}{p}}$$

There is also a "philosophical" reason:

$$\frac{1}{\text{probability of success}}$$

$p = 1/n$

average of n trials

$$\boxed{\mathrm{var}(N_p) = \frac{1-p}{p^2}} \qquad \boxed{\sigma(N_p) = \frac{\sqrt{1-p}}{p}}$$

The negative binomial distribution for general r:

Keep performing independent Bernoulli trials, each with p = probability of success. Counting the number of trials needed to succeed r times $=: N_{r,p}$

$N_{r,p}$ = the distribution of r independent random variables each of which has distribution N_p

$$\boxed{E(N_{r,p}) = \frac{r}{p}} \qquad \boxed{\operatorname{var}(N_{r,p}) = \frac{r(1-p)}{p^2}}$$

The probability mass function of the negative binomial distribution:

$$\text{Values} = \{r, r+1, r+2, \dots\} \;(\cup \{\infty\})$$

$p = 0$

$$P(N_{r,p} = n) = r \text{ successes}$$

includes the last trial performed!

$$\boxed{P(N_{r,p} = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}}$$

Example: The probability of a clover being 4-leaf is 10^{-4}.

(a) What is the probability it will take me exactly 1000 trials to find 3 four-leaf clovers? (Not a good question - answer extremely low).

(b) How many trials on average do I have to perform to find 3 four-leaf clovers? What is the standard deviation?

Solution: $P(N_{r,p} = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}$

$$r = 3, \; n = 1000, \; p = \frac{1}{10000} = 10^{-4}$$

(a) Answer: $\underline{\underline{\binom{999}{2} \cdot 10^{-12} \cdot (1-10^{-4})^{997}}}$ $\approx 10^3$

(b) $E(N_{r,p}) = \frac{r}{p}$

$r = 3$

$p = 10^{-4}$

$$E(N_{3,10^{-4}}) = \underline{\underline{3 \cdot 10^{4}}}$$

$$\text{var}(N_{3,10^{-4}}) = \frac{3 \cdot (1-10^{-4})}{10^{-8}} = 10^{8} \cdot 3 \cdot (1-10^{-4})$$

$$\sigma(N_{3,10^{-4}}) = \underline{\underline{10^{4} \cdot \sqrt{3}\sqrt{1-10^{-4}}}} \approx \underline{\underline{16{,}900}}$$

≈ 1

$\sqrt{3} \approx 1{,}69$

HW 3 I am casting a standard cubical die.

a What is the probability it will take exactly 10 trials for 6 to come up 4 times?

b What is the average number of trials it will take for 6 to come up 4 times?

c What is the standard deviation?

d Name the distribution (with numerical values of all the parameters).

Exam 2: From conditional probability through discrete random variables