Discrete Random Variables

Statistics 251: Statistical Methods

Module 4

Updated 2021

Terminology

random variable: describes the outcomes of a statistical experiment; a valued function

discrete variable: values that are countable; whole number values (i.e. chapters in a book, number of classes you can take)

Rules:

(1) $0 \leq P(A) \leq 1$, probabilities must be between 0 and 1 (0% and 100%)

(2) $\sum P(A_i) = 1 = S$, the sum of the probabilities for an experiment must sum to 1 (the sample space)
(3) $P(A') = 1 - P(A)$, the complement rule

example random variable

Suppose Nancy has classes three days a week. She attends classes three days a week 80% of the time, two days 15% of the time, one day 4% of the time, and no days 1% of the time. Suppose one week is randomly selected.

Table 9: Nancy's classes

X=# classes attended/week	0	1	2	3
$P(x_i)$	0.01	0.04	0.15	0.8

Figure 1: Nancy

Rules of expectation (mean, variance, standard deviation)

Since not all values of a random variable have the same probability, to calculate the mean, we have to approach it in a slightly different way. The mean is called an Expected Value (E(X)). It is a weighted mean (weighted average); meaning some values have a greater chance of happening than others.

$$E(X) = \sum x * p(x_i)$$

$$V(X) = \sum (x_i - E(X))^2 p(x_i)$$

$$SD(X) = \sqrt{V(X)}$$

Nancy's classes

Find the mean, variance, and standard deviation of the number of days in a week Nancy attends classes. Examples like Nancy's classes, flipping two coins, etc. are refered to as "generic" discrete probability distributions. There are special distributions that are "named" and used often to model situations.

$E(X) = \sum xp(x) = 0(0.01) + 1(0.04) + 2(0.15) + 3(0.8) = 2.74$ days on average

$V(X) = \sum(x - E(X))^2 p(x) = (0 - 2.74)^2(0.01) + (1 - 2.74)^2(0.04) + (2 - 2.74)^2(0.15) + (3 - 2.74)^2(0.8)$ $= 0.3324$

$SD(X) = \sqrt{\sigma^2} = \sqrt{0.3324} = 0.576541$

(1) Binomial distribution

There are three assumptions for an experiment to have a binomial distribution
(a) n, the number of trials, are fixed
(b) Only two outcomes are possible: success with probability p, and failure with probability $q = 1 - p$
(c) The n trials are indpendent

The probability is the probability of x successes out of n trials; any experiment that meets assumptions 2 and 3 when $n = 1$ is called a *Bernoulli Trial*, a binomial experiment happens when the number of successes is counted in one or more Bernoulli Trials.

Shorthand notation: $X \sim bin(n, p)$ or $X \sim B(n, p)$

Binomial formulas I

$$P(X = x) = \binom{n}{x} p^x q^{n-x} \quad x = 0, 1, \cdots, n$$

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

is a combination and reads as "n choose x." It is the number of ways that x things can be chosen from n trials (number of ways to get a sum of 7 when rolling two 6-sided dice, as an example).

Note that the exponents of p and q must sum to n.

Binomial formulas II

$$EX = np \quad VX = npq \quad SDX = \sqrt{npq}$$

You are allowed to use your calculator to do as much of the calculation as you want. There are instructions on how to use the command if you have a TI graphing calculator on page 254. Most scientific calculators have a ${}_nC_r$ command (same as "n choose x") for combinations

Binomial example

A trainer is teaching a dolphin to do tricks. The probability that the dolphin successfully performs the trick is 35%, and the probability that the dolphin does not successfully perform the trick is 65%. Out of 20 attempts, find the probability that the dolphin succeeds 12 times, succeeds *at most* 15 times.

$$X \sim bin(20, 0.35)$$

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

$$q = 1 - p = 1 - 0.35 = 0.65$$

Binomial example

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

$$P(X = 12) = \binom{20}{12} (0.35^{12})(0.65^8)$$

$$= 125970(0.35^{12})(0.65^8) = 0.013564$$

Binomial example

$$P(X \leq 15) = \sum_{0}^{15} bin(0 : 15, 20, 0.35)$$

$$= P(0) + P(1) + \cdots + P(15) = 1 - P(X > 15) = 1 - P(X \geq 16)$$

$$= 1 - [P(16) + P(17) + \cdots + P(20)]$$

$= 1 - [4845(0.35^{16})(0.65^4) + 1140(0.35^{17})(0.65^3)$
$+190(0.35^{12})(0.65^2) + 125970(0.35^{19})(0.65^1)$
$+125970(0.35^{20})(0.65^0)] = 1 - 0.00005 = 0.99995$

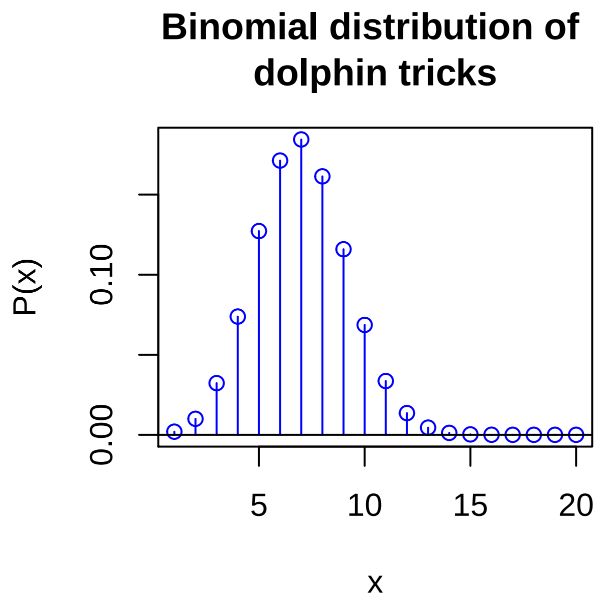
Binomial EX, VX, SDX

$$EX = np = 20(0.65) = 7$$

$$VX = npq = 20(0.35)(0.65) = 4.55$$

$$SDX = \sqrt{npq} = \sqrt{4.55} = 2.133073$$

Binomial Graph



(2) Poisson distribution

There are three assumptions for an experiment to have a Poisson distribution, and us used in modelling rare events

(a) The occurrence of one event does not affect the probability that a second event will occur. That is, events occur independently.

(b) The rate at which events occur is constant. The rate cannot be higher in some intervals and lower in other intervals.

(c) Two events cannot occur at exactly the same instant; instead, at each very small sub-interval exactly one event either occurs or does not occur.

The probability of the number of events happening within a specified period; the number of times an event
One of the things this distribution was used for was to model the number of horse kicks Prussian soldiers received (seriously, go look it up!)

Poisson logistics

Poisson is pronounced as "pwa-so(n)". If you have seen The Little Mermaid, recall the song that the chef sings when Sebastian is running from him in the kitchen Little Mermaid: Les Poissons

$$\text{Shorthand notation: } X \sim pois(\mu) \ or \ X \sim P(\mu)$$

μ is the average or a rate (use $\mu = np$, EX from binomial)
You are allowed to use your calculator to do as much of the calculation as you want. There are instructions on how to use the command if you have a TI graphing calculator on p. 267.

Poisson formulas

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!} \quad x = 0, 1, \cdots, \infty$$

$$EX = \mu$$

$$VX = \mu$$

$$SDX = \sqrt{\mu}$$

Poisson example

Consider an experiment that consists of counting the number of α-particles given off in a 1-second time interval by 1 gram of radioactive material. If the average number of α-particles given off is 3.2, what is the probability of exactly 2 α-particles given off in the next 1-second interval? What is the probability that no α-particles are given off in the next 1-second interval? More than 2 α-particles?

$$X \sim pois(3.2) \; or \; X \sim P(3.2)$$

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!}$$

Poisson example continued

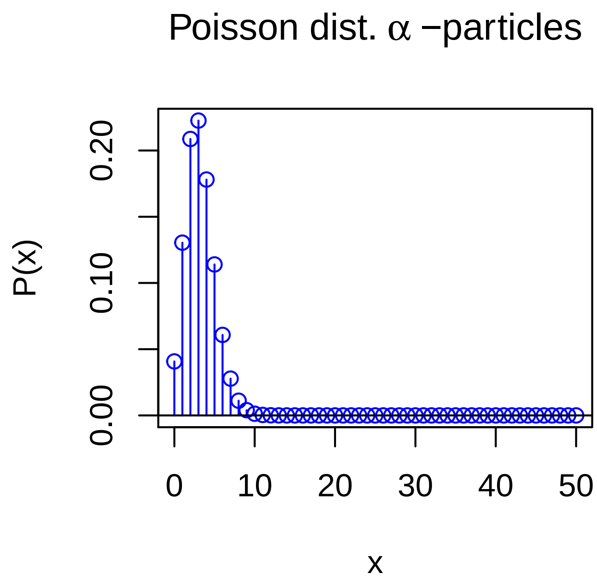
$$P(X = 2) = \frac{e^{-3.2}3.2^2}{2!} = 0.208702$$

$$P(X = 0) = \frac{e^{-3.2}3.2^0}{0!} = 0.040762$$

$$P(X > 2) = 1 - P(X \leq 2) = 1 - [P(2) + P(1) + P(0)]$$

$$= 1 - \left(\frac{e^{-3.2}3.2^2}{2!} + \frac{e^{-3.2}3.2^1}{1!} + \frac{e^{-3.2}3.2^0}{0!}\right)$$
$$= 1 - 0.379904 = 0.620096$$

Poisson graph



Poisson EX, VX, SDX

$$EX = \mu = 3.2$$

$$VX = \mu = 3.2$$

$$SDX = \sqrt{\mu} = \sqrt{3.2} = 1.788854$$