Prob/Stat/Discrete Math **Unit 5: Discrete Probability Distributions**

5.1 Probability Distributions

Section 5.1 Objectives:

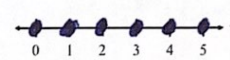
- Distinguish between discrete random variables and continuous random variables
- Construct a discrete probability distribution and its graph
- Determine if a distribution is a probability distribution
- Find the mean, variance, and standard deviation of a discrete probability distribution
- Find the expected value of a discrete probability distribution

Random Variables:

- Represents a numerical value associated with each outcome of a probability distribution.
- Denoted by ***x***
- Examples
  - x = Number of sales calls a salesperson makes in one day.
  - x = Hours spent on sales calls in one day.

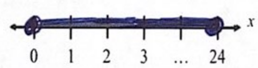
element of chance or randomness

- **Discrete Random Variable**
  - Has a finite or countable number of possible outcomes that can be listed.
  - Example
    - x = Number of sales calls a salesperson makes in one day.



usually whole numbers

- **Continuous Random Variable**
  - Has an infinite number of possible outcomes, represented by an interval on the number line.
  - Example
    - x = Hours spent on sales calls in one day.



includes fractions & decimals

Example: Random Variables

Decide whether the random variable x is discrete or continuous.

1. x = The number of stocks in the Dow Jones Industrial Average that have share price increases on a given day.

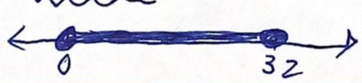
discrete ⇒ countable (whole #)

2. x = The volume of water in a 32-ounce container.

continuous ⇒ fractions & decimals make sense, so there is an infinite # of possible volumes

⇓

represent by an interval from 0 to 32



Discrete probability distribution:

- Lists each possible value the random variable can assume, together with its probability.
- Must satisfy the following conditions:
  - The probability of each value of the discrete random variable is between 0 and 1, inclusive.
    $0 \le P(x) \le 1$
  - The sum of all the probabilities is 1.
    $\Sigma P(x) = 1$

Constructing a Discrete Probability Distribution

Let x be a discrete random variable with possible outcomes $x_1, x_2, \ldots, x_n$.

1. Make a frequency distribution for the possible outcomes. frequency table
2. Find the sum of the frequencies. Σf
3. Find the probability of each possible outcome by dividing its frequency by the sum of the frequencies.
4. Check that each probability is between 0 and 1 and that the sum is 1.

Example: Constructing a Discrete Probability Distribution

An industrial psychologist administered a personality inventory test for passive-aggressive traits to 150 employees. Individuals were given a score from 1 to 5, where 1 was extremely passive and 5 extremely aggressive. A score of 3 indicated neither trait. Construct a probability distribution for the random variable x. Then graph the distribution using a histogram.

Score, x	Frequency, f	Probability $P(x)$
1	24	$P(1) = 24/150 = 0.16$
2	33	$P(2) = 33/150 = 0.22$
3	42	$P(3) = 42/150 = 0.28$
4	30	$P(4) = 30/150 = 0.20$
5	21	$P(5) = 21/150 = 0.14$

$\Sigma f = 150$ $\Sigma P(x) = 1.00$

same as the relative frequency column in the last chapter

Solution: Constructing a Discrete Probability Distribution

- Divide the frequency of each score by the total number of individuals in the study to find the probability for each value of the random variable.

→

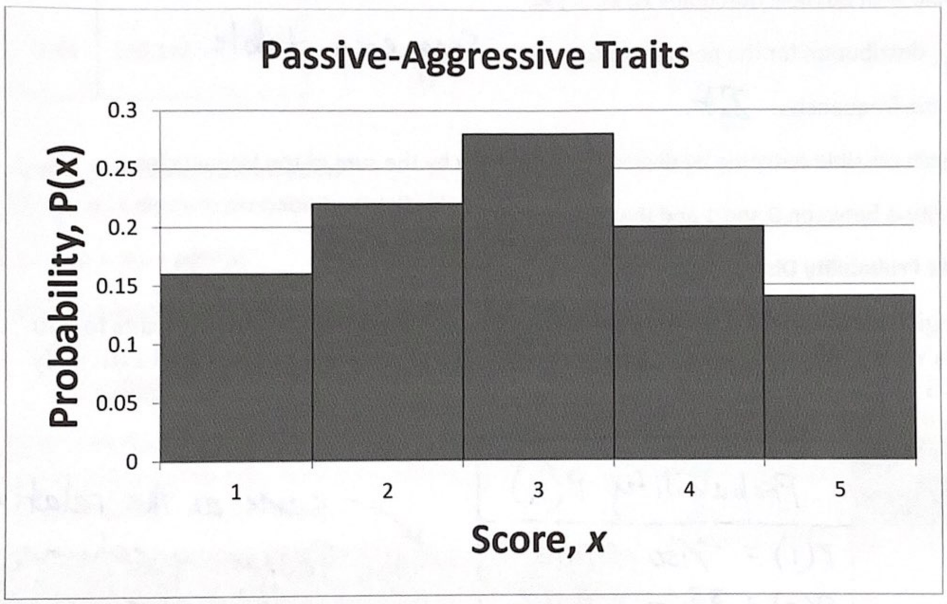
- **Discrete probability distribution:**

x	1	2	3	4	5
xP(x)	0.16	0.22	0.28	0.20	0.14

This is a valid discrete probability distribution since

✓ 1. Each probability is between 0 and 1, inclusive, $0 \le P(x) \le 1$.

✓ 2. The sum of the probabilities equals 1, $\Sigma P(x) = 0.16 + 0.22 + 0.28 + 0.20 + 0.14 = 1$.



Width of each bar is always 1 in a discrete prob. distribution histogram (Different from relative frequency histogram)

Because the width of each bar is one, the area of each bar is equal to the probability of a particular outcome.

Mean

Mean of a discrete probability distribution $\mu = \Sigma [x \cdot P(x)]$

Recall: "expected value" from earlier unit!

$\mu = \Sigma xP(x)$

- **Each value of *x* is** multiplied **by its corresponding probability and the products are added.**

Example: Finding the Mean $\mu = \Sigma[x \cdot P(x)]$ "Theoritcal avg" – sometimes not a possible outcome.

The probability distribution for the personality inventory test for passive-aggressive traits is given. Find the mean.

→ When have we done this before?

→ $\boxed{\text{Expected Value}}$

↳ (only made sense for "in the long run")

Solution:

x	P(x)	xP(x)
1	0.16	1(0.16) = .16
2	0.22	2(0.22) = .44
3	0.28	3(0.28) = .84
4	0.20	4(0.20) = .80
5	0.14	5(0.14) = .70

$\Sigma x \cdot P(x) = 2.94$ ← MEAN

(Look back at graph – centered @ 3 but just a little higher to left)

Variance and Standard Deviation:

Variance of a discrete probability distribution

- $\sigma^2 = \Sigma[(x-\mu)^2 P(x)]$

Standard deviation of a discrete probability distribution

- $\sigma = \sqrt{\sigma^2} = \sqrt{\Sigma[(x-\mu)^2 P(x)]}$

before: $\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\Sigma(x-\mu)^2}{N}}$

Example: Finding the Variance and Standard Deviation:

The probability distribution for the personality inventory test for passive-aggressive traits is given. Find the variance and standard deviation. (μ = 2.94) (above)

x	P(x)	x − μ	(x − μ)²	(x − μ)²P(x)
1	0.16	1−2.94 = −1.94	3.764	(3.764)(.16) = .602
2	0.22	2−2.94 = −0.94	0.884	.194
3	0.28	3−2.94 = 0.06	0.004	.001
4	0.20	4−2.94 = 1.06	1.124	.225
5	0.14	5−2.94 = 2.06	4.244	.594

Variance: $\sigma^2 = \Sigma(x-\mu)^2 P(x) = 1.616$

≈ 1.6

Standard Deviation: $\sqrt{1.616}$

≈ 1.3

$\Sigma[(x-\mu)^2 P(x)] = 1.616$ ← VARIANCE

Expected Value:

Expected value of a discrete random variable → what you would expect, on the avg, in thousands of trials

- Equal to the Mean of the random variable.
- $E(x) = \mu = \Sigma x P(x)$ = $\Sigma x \cdot P(x)$

Example: Finding an Expected Value

At a raffle, 1500 tickets are sold at $2 each for four prizes of $500, $250, $150, and $75. You buy one ticket. What is the expected value of your gain?

- **Solution: Finding an Expected Value**
  To find the gain for each prize, subtract the price of the ticket from the prize:
  - Your gain for the $500 prize is 498
  - Your gain for the $250 prize is 248
  - Your gain for the $150 prize is 148
  - Your gain for the $75 prize is 73
- If you do not win a prize, your gain is -2
- **Solution: Finding an Expected Value**
  Probability distribution for the possible gains (outcomes)

Gain, x	498	248	148	73	-2
P(x)	$\frac{1}{1500}$	$\frac{1}{1500}$	$\frac{1}{1500}$	$\frac{1}{1500}$	$\frac{1496}{1500}$

← all the rest, after 4 prizes

You can expect to lose an average of $ 1.35 for each ticket you buy.

$$498 \cdot \frac{1}{1500} + 248 \cdot \frac{1}{1500} + 148 \cdot \frac{1}{1500} + 73 \cdot \frac{1}{1500} + -2 \cdot \frac{1496}{1500}$$

$$= \frac{498}{1500} + \frac{248}{1500} + \frac{148}{1500} + \frac{73}{1500} - \frac{2992}{1500}$$

$$= \frac{-2025}{1500} = \boxed{-1.35}$$

Section 5.2: Binomial Distributions

Section 5.2 Objectives:

- Determine if a probability experiment is a binomial experiment
- Find binomial probabilities using the binomial probability formula
- Find binomial probabilities using technology and a binomial table
- Graph a binomial distribution
- Find the mean, variance, and standard deviation of a binomial probability distribution

Binomial Experiments:

1. The experiment is repeated for a fixed n number of trials, where each trial is independent of other trials.
2. There are only 2 possible outcomes of interest for each trial. The outcomes can be classified as a success (*S*) or as a failure (*F*).
3. The probability of a success *P*(*S*) is the same for each trial. $p = P(S)$, $q = P(F)$ failure
4. The random variable *x* counts the number of successful trials.

$x = 0, 1, 2, 3, \ldots n$

Notation for Binomial Experiments

Symbol	*Description*
n	The number of times a trial is repeated
$\boxed{p = P(s)}$	The probability of success in a single trial
$\boxed{q = P(F)}$	The probability of failure in a single trial ($q = 1 - p$)
X	The random variable represents a count of the number of successes in *n* trials: $x = 0, 1, 2, 3, \ldots, n$.

Ex. If $P(S) = p = .2$
Then $P(F) = q = 1 - .2 = .8$

no successes

all successes

Example: Binomial Experiments

Decide whether the experiment is a binomial experiment. If it is, specify the values of *n*, *p*, and *q*, and list the possible values of the random variable *x*.

① number of trials, $n = 3$

2. A jar contains five red marbles, nine blue marbles, and six green marbles. You randomly select three marbles from the jar, *without replacement*. The random variable represents the number of red marbles.

② $x = \#$ red → so red is Success

③ $P(S) = \frac{5}{20}$ first trial

$\frac{4}{19}$ 2nd trial

not indep. events ① failed

So $P(S)$ NOT same for each trial ⇒ NOT Binomial Exper.

Binomial Probability Formula

- The probability of exactly *x* successes in *n* trials is

$$\boxed{P(x) = {}_nC_x p^x q^{n-x} = \frac{n!}{(n-x)!x!} p^x q^{n-x}}$$

- n = number of trials
- p = probability of success
- $q = 1 - p$ probability of failure
- x = number of successes in n trials

Before: X counts the # of successful trials

SAME

Example: Finding Binomial Probabilities

$p = P(S) = .75$

Microfracture knee surgery has a 75% chance of success on patients with degenerative knees. The surgery is performed on three patients. Find the probability of the surgery being successful on exactly two patients.

$n = 3$

Solution: Finding Binomial Probabilities

Method 1: Draw a tree diagram and use the Multiplication Rule

1st Surgery	2nd Surgery	3rd Surgery	Outcome	Number of Successes	Probability (and and)
S	S	S	SSS	3	$\frac{3}{4}\cdot\frac{3}{4}\cdot\frac{3}{4} = \frac{27}{64}$
S	S	F	SSF	2	$\frac{3}{4}\cdot\frac{3}{4}\cdot\frac{1}{4} = \frac{9}{64}$
S	F	S	SFS	2	$\frac{3}{4}\cdot\frac{1}{4}\cdot\frac{3}{4} = \frac{9}{64}$
S	F	F	SFF	1	$\frac{3}{4}\cdot\frac{1}{4}\cdot\frac{1}{4} = \frac{3}{64}$
F	S	S	FSS	2	$\frac{1}{4}\cdot\frac{3}{4}\cdot\frac{3}{4} = \frac{9}{64}$
F	S	F	FSF	1	$\frac{1}{4}\cdot\frac{3}{4}\cdot\frac{1}{4} = \frac{3}{64}$
F	F	S	FFS	1	$\frac{1}{4}\cdot\frac{1}{4}\cdot\frac{3}{4} = \frac{3}{64}$
F	F	F	FFF	0	$\frac{1}{4}\cdot\frac{1}{4}\cdot\frac{1}{4} = \frac{1}{64}$

So, outcomes are SSF OR SFS or FSS

Prob. is $\frac{9}{64} + \frac{9}{64} + \frac{9}{64} = 3\left(\frac{9}{64}\right) = \frac{27}{64} \approx 0.422$

"x" counts the # of successes & we're looking for 2

ethod 2: Binomial Probability Formula

$n = 3$ trials $p = .75$ ↳ P(S) $q = .25$ ↳ $P(F) = 1 - .75$ $x = 2$

$$P(x \text{ successful surgeries}) = \frac{n!}{(n-x)!\,x!} \cdot p^{x} \cdot q^{n-x}$$

$$\rightarrow P(2 \text{ successful surgeries}) = \frac{3!}{(3-2)!\,2!} \cdot \left(\frac{3}{4}\right)^2 \cdot \left(\frac{1}{4}\right)^1$$

$$= \frac{3 \cdot \cancel{2 \cdot 1}}{1!\,\cancel{2 \cdot 1}} \cdot \frac{9}{16} \cdot \frac{1}{4} = 3 \cdot \frac{9}{64} = \frac{27}{64} \approx \boxed{.422}$$

P(2 successes)

$$= {}_nC_x\, p^x q^{n-x}$$
$$= {}_3C_2\, (.75)^2 (.25)^{3-2}$$
$$= 3(.5625)(.25)^1$$
$$= 0.421875 \approx \boxed{0.422}$$

Binomial Probability Distribution → Table

- List the possible values of *x* with the corresponding probability of each.
- Example: Binomial probability distribution for Microfacture knee surgery: $n = 3$, $p = ¾$

Note: $0! = 1$

$$\frac{3!}{(3-1)!\,1!}(.75)^1(.25)^2 = \frac{3!}{2!}(.75)(.25)^2 = 0.140625$$

Use binomial probability formula to find probabilities. $x = 2$ is done

x	0 ($\frac{1}{64}$ ✓)	1 ($3 \cdot \frac{3}{64}$ ✓)	2 (DONE)	3 ($\frac{27}{64}$ ✓)
$P(x)$	.016	0.141	0.422	0.422

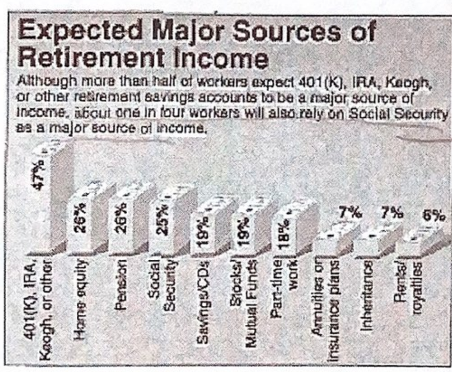
$$\frac{3!}{(3-0)!\,0!}(.75)^0(.25)^3 = 1 \cdot 1 \cdot (.25)^3 = .015625$$

$$\frac{3!}{(3-3)!\,3!}(.75)^3(.25)^0 = 1\,(.75)^3 \cdot 1 = .421875$$

Do this one!

Example: Constructing a Binomial Distribution:

In a survey, workers in the U.S. were asked to name their expected sources of retirement income. Seven workers who participated in the survey are randomly selected and asked whether they expect to rely on Social Security for retirement income. Create a binomial probability distribution for the number of workers who respond yes. $n = 7$ Success: Yes



(Source: The Gallup Organization)

"about one in four workers will also rely on Soc.Sec. as a major source of income"

$p = P(S) = .25$

$q = P(F) = .75$

$P(x) = {}_nC_x \, p^x \cdot q^{n-x}$ [illegible]

[illegible]

Solution: Constructing a Binomial Distribution

- [illegible]
- [illegible]

[illegible]

x	$P(x)$
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]

$\Sigma P(x) = 1$

[illegible]

Example: Finding Binomial Probabilities

[illegible]

Solution:

- [illegible]
- [illegible]
- [illegible]

Solution: Finding Binomial Probabilities

[illegible]

Example: Finding Binomial Probabilities Using Technology

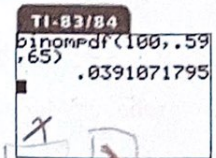
$p = P(S) = .59$

The results of a recent survey indicate that when grilling, 59% of households in the United States use a gas grill. If you randomly select 100 households, what is the probability that exactly 65 households use a gas grill? Use a technology tool to find the probability. *(Source: Greenfield Online for Weber-Stephens Products Company)*

$n=100$ $x=65$

Solution:

- Binomial with n = 100, p = 0.59, x = 65



Solution: Finding Binomial Probabilities Using Technology:

2nd DISTR VARS ↓ menu A: binompdf(n 100 , p 0.59 , x 65) ≈ 0.04

From the display, you can see that the probability that exactly 65 households use a gas grill is about 0.04.

Example: Finding Binomial Probabilities Using a Table

About thirty percent of working adults spend less than 15 minutes each way commuting to their jobs. You randomly select six working adults. What is the probability that exactly three of them spend less than 15 minutes each way commuting to work? Use a table to find the probability. *(Source: U.S. Census Bureau)*

Solution:

- Binomial with n = 6, p = .30, x = 3

Solution: Finding Binomial Probabilities Using a Table

- A portion of Table 2 is shown

n	x	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60
2	0	.980	.902	.810	.723	.640	.563	.490	.423	.360	.303	.250	.203	.160
	1	.020	.095	.180	.255	.320	.375	.420	.455	.480	.495	.500	.495	.480
	2	.000	.002	.010	.023	.040	.063	.090	.123	.160	.203	.250	.303	.360
3	0	.970	.857	.729	.614	.512	.422	.343	.275	.216	.166	.125	.091	.064
	1	.029	.135	.243	.325	.384	.422	.441	.444	.432	.408	.375	.334	.288
	2	.000	.007	.027	.057	.096	.141	.189	.239	.288	.334	.375	.408	.432
	3	.000	.000	.001	.003	.008	.016	.027	.043	.064	.091	.125	.166	.216
6	0	.941	.735	.531	.377	.262	.178	.118	.075	.047	.028	.016	.008	.004
	1	.057	.232	.354	.399	.393	.356	.303	.244	.187	.136	.094	.061	.037
	2	.001	.031	.098	.176	.246	.297	.324	.328	.311	.278	.234	.186	.138
	3	.000	.002	.015	.042	.082	.132	.185	.236	.276	.303	.312	.303	.276
	4	.000	.000	.001	.006	.015	.033	.060	.095	.138	.186	.234	.278	.311
	5	.000	.000	.000	.000	.002	.004	.010	.020	.037	.061	.094	.136	.187
	6	.000	.000	.000	.000	.000	.000	.001	.002	.004	.008	.016	.028	.047

$n=6$ → row n = 6; $x=3$ → row x = 3

The probability that exactly three of the six workers spend less than 15 minutes each way commuting to work is 0.185.

10 min

$$P(x) = {}_nC_x \cdot p^x \cdot q^{n-x} = \frac{n!}{(n-x)!\,x!} \cdot p^x \cdot q^x$$

Example: Graphing a Binomial Distribution

Fifty-nine percent of households in the U.S. subscribe to cable TV. You randomly select six households and ask each if they subscribe to cable TV. Construct a probability distribution for the random variable *x*. Then graph the distribution. *(Source: Kagan Research, LLC)*

3 methods

Solution:

- $n =$ 6, $p =$ .59, $q =$ .41
- Find the probability for each value of x

$$P(x=0) = \boxed{{}_6C_0}(.59)^0(.41)^6 = 1 \cdot 1 \cdot .00475 \approx .005$$

$$P(x=1) = \boxed{\frac{6!}{(6-1)!\,1!}}(.59)^1(.41)^5 = \frac{6 \cdot 5!}{5!\,1!}(.59)(.41)^5 = 6\,(.59)(.41)^5 \approx .041$$

Solution: Graphing a Binomial Distribution

x	0	1	2	3	4	5	6
P(x)	0.005	0.041	0.148	0.283	0.306	0.176	0.042

$P(x=2) =$

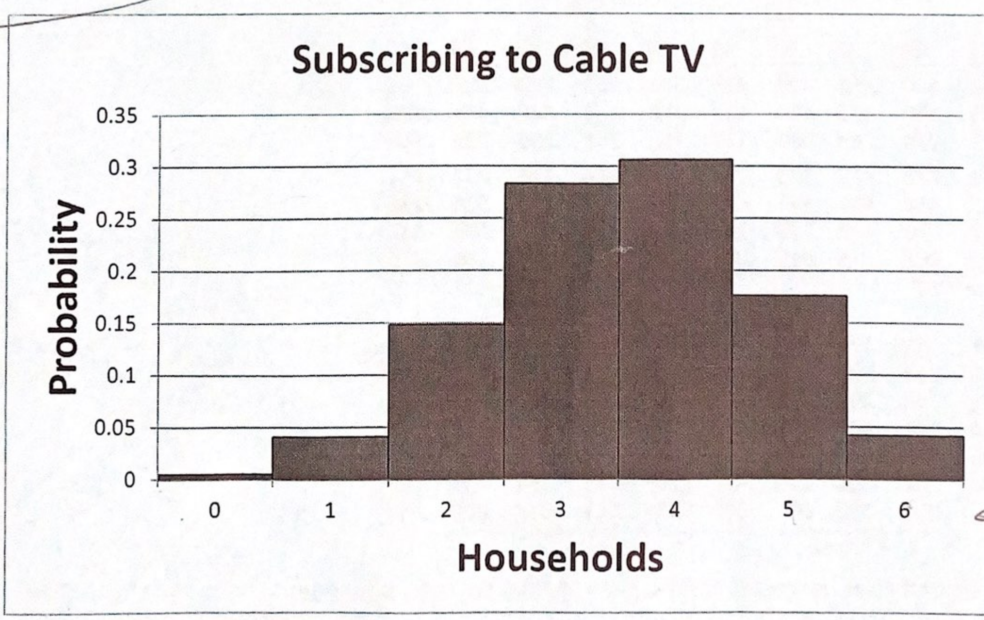
2nd DIST
binompdf(n, p, x)

binompdf(6, .59, 2)

≈ 0.148

WRITE to show work!

Histogram:



quantitative

Mean, Variance, and Standard Deviation:

- **Mean:** $\mu = np$ $\quad \mu = n \cdot p$
- **Variance:** $\sigma^2 = npq$ $\quad Var = n \cdot p \cdot q$
- **Standard Deviation:** $\boxed{\sigma = \sqrt{npq}}$

Example: Finding the Mean, Variance, and Standard Deviation

$p = .56$

In Pittsburgh, Pennsylvania, about 56% of the days in a year are cloudy. Find the mean, variance, and standard deviation for the number of cloudy days during the month of June. Interpret the results and determine any unusual values. *(Source: National Climatic Data Center)*

→ 30 days

Solution: $n =$ 30, $p =$ .56, $q =$ .44

Mean: $\mu = np = 30(.56) = 16.8$

Variance: $\sigma^2 = npq = 30(.56)(.44) \approx 7.4$

Standard Deviation: $\sqrt{\sigma^2} = \sigma$ or $\sqrt{npq} = \sqrt{7.4} \approx 2.7$

Solution: Finding the Mean, Variance, and Standard Deviation

$\mu =$ 16.8 $\quad \sigma^2 \approx$ 7.4 $\quad \sigma \approx$ 2.7

- On average, there are 16.8 cloudy days during the month of June.
- The standard deviation is about 2.7 days.
- Values that are more than two standard deviations from the mean are considered unusual.

to the left ▪ $16.8 - 2(2.7) = 11.4$, A June with 11 cloudy days would be unusual.

to the right ▪ $16.8 + 2(2.7) = 22.2$, A June with 23 cloudy days would also be unusual.