Binomial and Normal Distribution

York College & University Centre

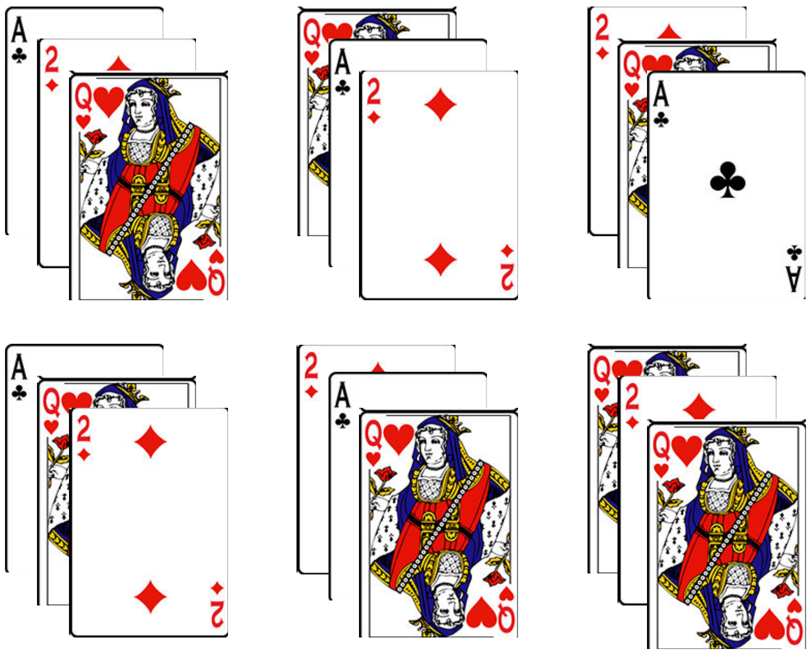
Factorials

- A factorial is a symbol to simplify equations
- n! is equal to the product of the integers from 1 to n
- For example, 4! = $4 * 3 * 2 * 1$ = 24
- 5! = $5 * 4 * 3 * 2 * 1$ = 120

n! = n
0 = 1
1 = 1
2 = 2
3 = 6
4 = 24
5 = 120
6 = 720
7 = 5.040
8 = 40.320
9 = 362.880
10 = 3.628.800
11 = 39.916.800
12 = 479.001.600

Factorials

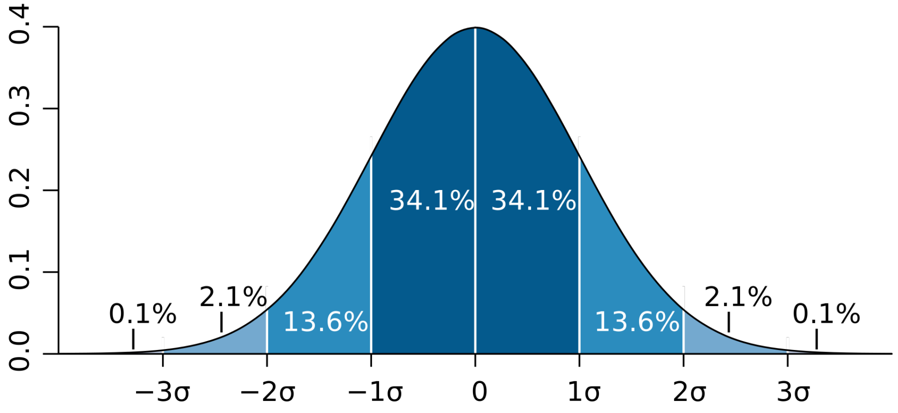
- The best way to think about factorials is the ways of arranging objects.
- If we have 3 playing cards, then there are 6 different ways we can lay those cards out, this is equal to 3! Which is 6
- This is why 0! is 1 because there is only 1 way to arrange no cards.



Binomial & Normal Distribution

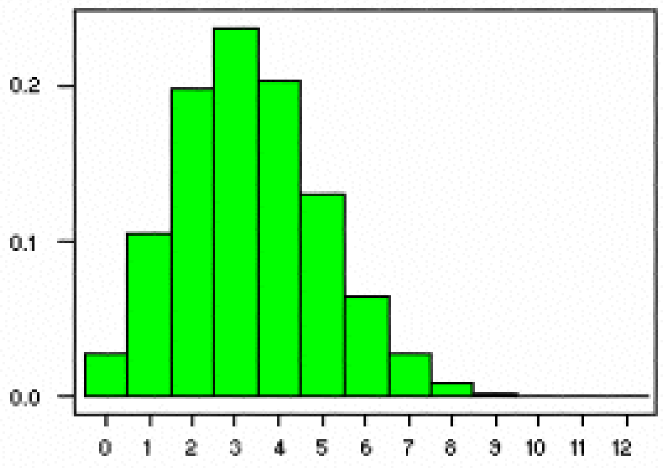
- Binomial and normal distributions are both just methods used to model real-world randomness and uncertainty in a mathematically useful way.

- They are helpful in making predictions, understanding variability and decision making.



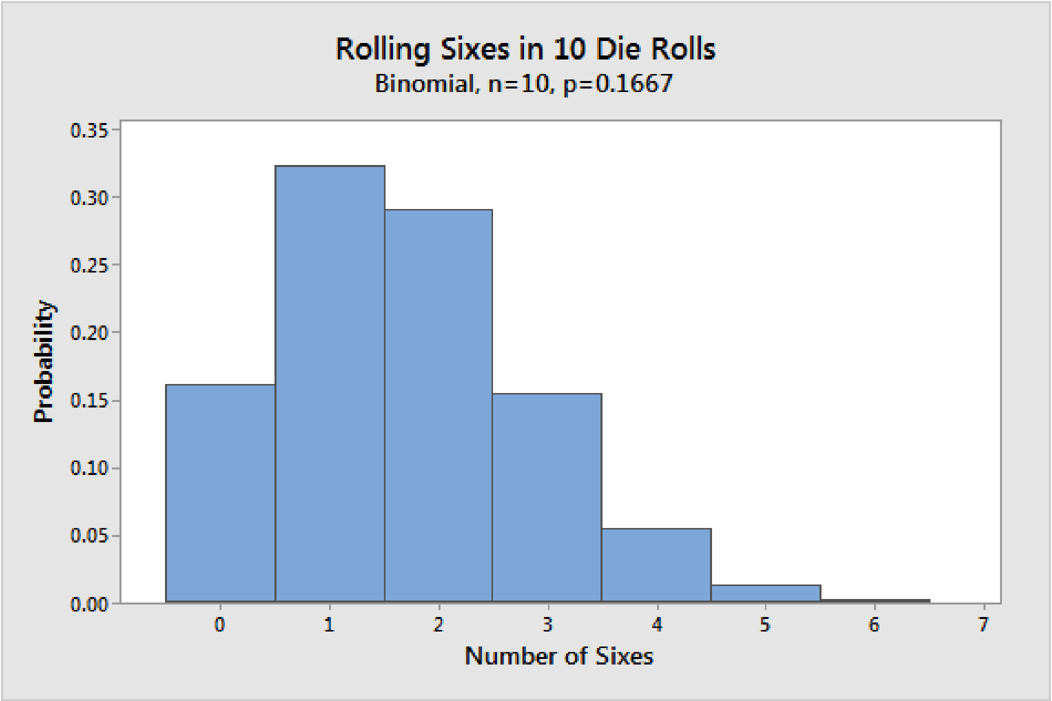
Normal

Binomial



Binomial Distribution

- A binomial distribution is a **discrete** probability distribution
- This means the values are countable and separate
- Models the probability of something happening in a set amount of trials
- Example: flipping a coin 10 times and counting how many times heads appears



Binomial Distribution Formula

- Follows the rule:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

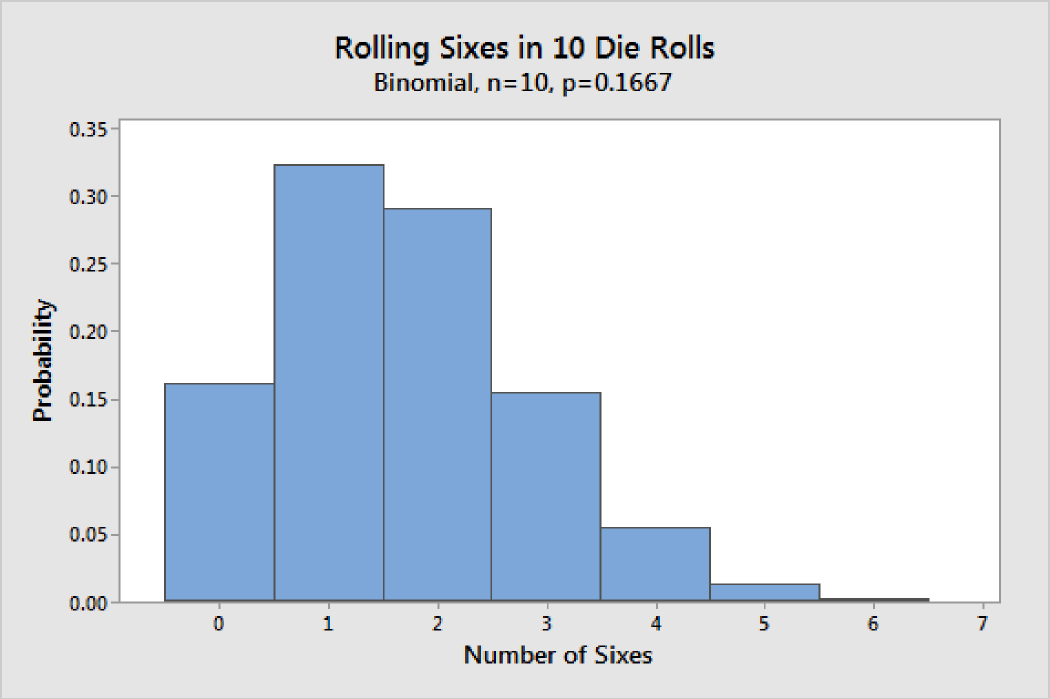
- Where:

k is the amount of successes

n is the number of trials

p is the probability of success in a single trial

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



Probability of success (p)

- We can work out the probability of success in an individual test using the equation:

$$p = \frac{Number\ of\ Success}{Total\ number\ of\ trials}$$

Example: if we flip a coin 10 times and it lands on heads 6 times the probability for heads is 6/10 = 0.6

Example of solving a binomial question

- A power company manufactures **circuit breakers,** each breaker has a **5% probability** of being faulty. If a technician tests 10 randomly selected breakers, what is the probability that exactly 2 of them are faulty?

Example of solving a binomial question

- Identify the given values:

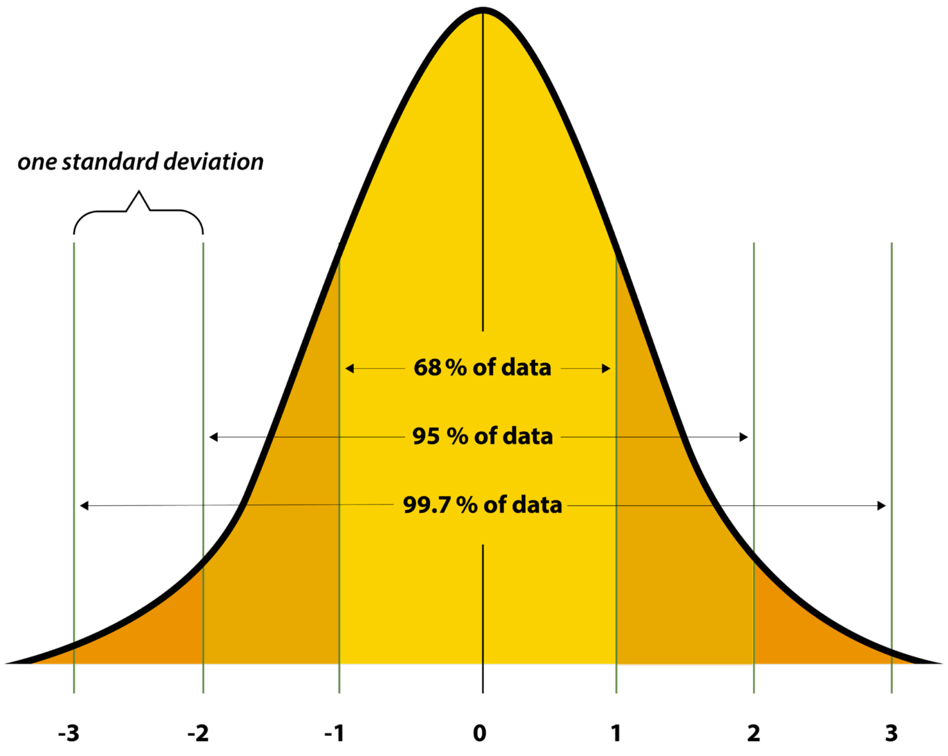
- Total trials: **n = 10**
- Probability of success: **p = 0.05**
- Desired successes: **k = 2**

Practical

- In pairs complete the Binomial Distribution Practical
- The sheets explain everything you need and what to do

Normal Distribution

- A binomial distribution is a continuous (Gaussian) probability distribution
- This means the values can be anything within a range
- Takes the form of a symmetrical bell-shaped curve defined by mean and standard deviation
- Example: exam scores will follow a normal distribution allowing universities to set grading curves



Z-score Equation

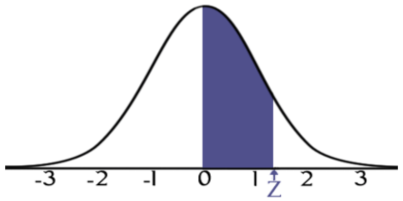
- The first formula to do with normal distribution is the z-score equation which is:

$$Z = \frac{X - \mu}{\sigma}$$

- Where:
  - X = the data point
  - μ = the mean
  - σ = the standard deviation

- Used when trying to find the probability of something happening in a normal distribution

- Uses the z-table to find the probability



STANDARD NORMAL TABLE (Z)

Entries in the table give the area under the curve between the mean and z standard deviations above the mean. For example, for $z = 1.25$ the area under the curve between the mean (0) and z is 0.3944.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0190	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2969	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3513	0.3554	0.3577	0.3529	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998

Reading a z-table

- Z-tables look complicated but they're easy
- All you do is append the column to the row
- example, for example for z=-2.51 we would go to the tenth row (-2.5) then the second column (.01)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681

Probability Density Formula

- The second equation to do with Normal distribution is the probability density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

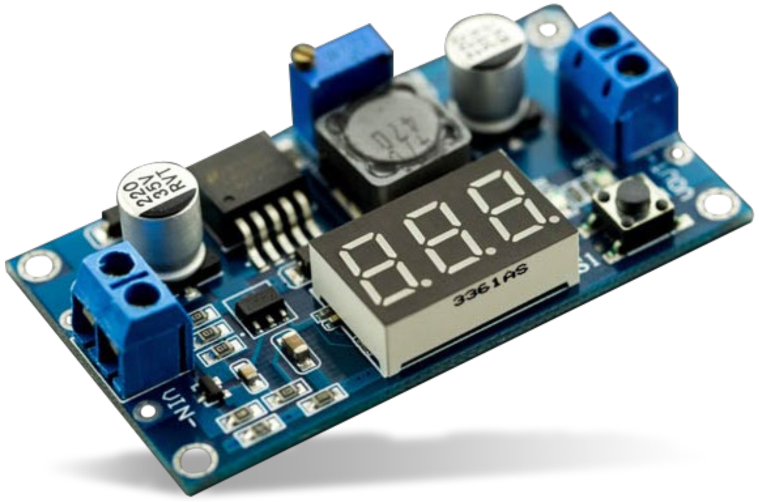
- Where:
  - x is the value we are trying to find (e.g a height of 175cm)
  - μ is the mean average around which all data is centered (the peak of the bell curve)
  - σ is the standard deviation (how far apart the data is)

- Used when determining how dense probability is at a single point, for calculation the relative likelihood of different values

- The density will not give us a probability, it will tell us how likely the height is in that region

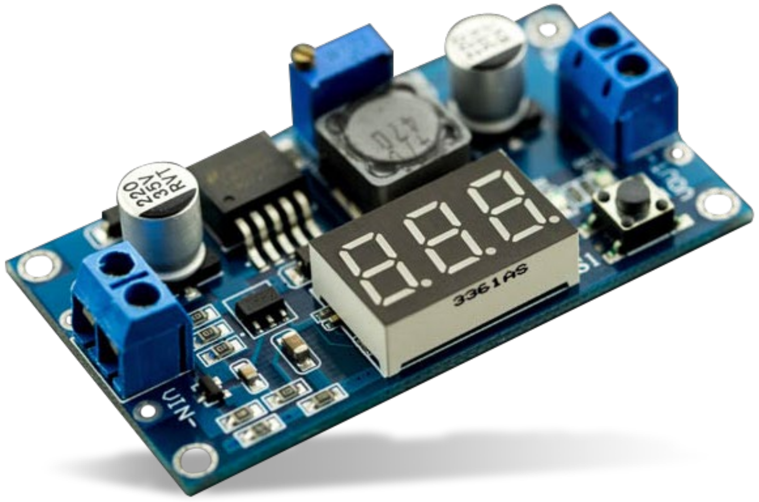
Example of solving a normal distribution (z-score)

- A component has a mean voltage output of 12v ($\mu = 12$)
- They are built with a specification standard deviation of 0.5v ($\sigma = 0.5$)
- A component will have to be thrown out if it has an output voltage **below 11v**, what is the probability that a randomly chosen power supply will trigger the alarm?



Example of solving a normal distribution (z-score)

- Step 1 is to define the problem:
  - $P(X < 11)$
- Step 2 is to put the values into the z-score equation
- $Z = \frac{X-\mu}{\sigma} = \frac{11-12}{0.5} = -2$



Example of solving a normal distribution (z-score)

- Step 3 is to find the probability from the z-table

- Z = -2.00

- Therefore probability = 0.0228

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
–3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
–3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
–3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
–3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
–3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
–2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
–2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
–2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
–2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
–2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
–2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
–2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
–2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
–2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
–2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
–1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233

Example of solving a normal distribution (Probability Density Function)

- Suppose the heights of students in a class are normally distributed with:
- Mean = 170cm
- Standard Deviation = 5cm
- What is the probability density of a student having a height of 172cm

Example of solving a normal distribution (Probability Density Function)

- Step 1: Write out the formula

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Step 2: Plug in values

$$f(172) = \frac{1}{5\sqrt{2\pi}} e^{-\frac{(172-170)^2}{2*5^2}}$$

Example of solving a normal distribution (Probability Density Function)

- Step 3: Simplify the equation

$$f(172) = \frac{1}{5\sqrt{2\pi}} e^{-\frac{(2)^2}{50}} = \frac{1}{5\sqrt{2\pi}} e^{-0.08}$$

- Step 4: Calculate components

$$f(172) = \frac{1}{5 * 2.5066} * 0.9231 = 0.0735$$

- Step 5: Interpret the results

The probability density at 172cm is 0.0735. This means that the relative likelihood of a student being exactly 172cm tall is 0.0735, bear in mind this is not the true probability, just an idea of how likely the height is in that ballpark

Your Turn

- **Question 1:** The heights of a population of adults follow a normal distribution with a **mean of 170 cm** and a **standard deviation of 7 cm**.

  a) Calculate the Z-score for a person with a height of 178 cm.

  b) Work out the probability of having the height 178cm.

  b) Using the PDF formula, calculate the probability density for a person with a height of 178 cm.

- **Question 2:** The test scores of a class follow a normal distribution with a **mean score of 75** and a **standard deviation of 8**.

  a) A student scored 65. Calculate the Z-score.

  b) Interpret the result: Is this score below average? How far is it from the mean in terms of standard deviations?