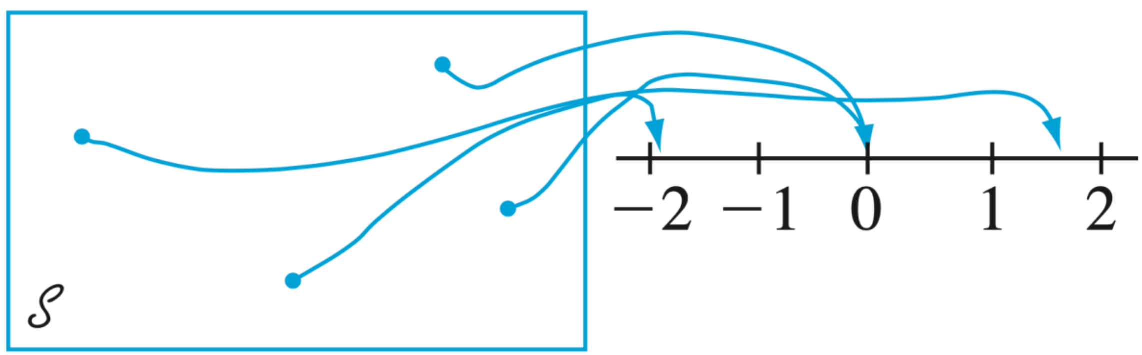
Probability Theory Review
Part 2

Overview

- Random Variables
- Expected Value
- Pairs of Discrete Random Variables
  – Conditional Probability
  – Bayes Rule
- Continuous Random Variables

Random Variables

- A ***Random Variable*** ***is a measurement on an outcome of a random experiment*** – denoted by r.v. X.



Probability and statistics for engineering and the sciences – Jav Devore

Types of Random Variables

Discrete versus ***Continuous*** ***random variable:*** an r.v. X is discrete if it can assume a finite or countably infinite number of values. An r.v. X is continuous if it can assume all values in an interval.

Examples

- Which of the following random variables are discrete and which are continuous?

- X = Number of houses sold by real estate developer per week?
- X = Number of heads in ten tosses of a coin?
- X = Weight of a child at birth?
- X = Time required to run 100 yards?

Probability Mass Function

A (discrete) random variable has an associated probability mass function (PMF), which gives the probability of each numerical value that the random variable can take.

The probability mass function (PMF) of a discrete random variable X, is denoted by

$$p(x) = P(X = x) = P(all\ \omega \in \Omega : X(\omega) = x)$$

Probability Mass Function Example

Let the experiment consist of two independent tosses of a fair coin, and let X be the number of heads obtained. What is PMF of X?

Probability Mass Function Example

Let the experiment consist of two independent tosses of a fair coin, and let X be the number of heads obtained. What is PMF of X?

$$p_X(x) = \begin{cases} 1/4 & \textit{if } x = 0 \textit{ or } x = 2 \\ 1/2 & \textit{if } x = 1 \\ 0 & \textit{otherwise} \end{cases}$$

Probability Distribution Example: X is the Sum of Two Dice

red	1	2	3	4	5	6

This sequence provides an example of a discrete random variable. Suppose that you have a red die which, when thrown, takes the numbers from 1 to 6 with equal probability.

Probability Distribution Example: X is the Sum of Two Dice

red / green	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

Suppose that you also have a green die that can take the numbers from 1 to 6 with equal probability.

Probability Distribution Example: *X* is the Sum of Two Dice

red green	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

We will define a random variable *X* as the sum of the numbers when the dice are thrown.

Probability Distribution Example: *X* is the Sum of Two Dice

red / green	1	2	3	4	5	6
1						
2						
3						
4						
5						
6				10		

For example, if the red die is 4 and the green one is 6, *X* is equal to 10.

Probability Distribution Example: X is the Sum of Two Dice

red / green	1	2	3	4	5	6
1						
2						
3						
4						
5		7				
6						

Similarly, if the red die is 2 and the green one is 5, *X* is equal to 7.

Probability Distribution Example: X is the Sum of Two Dice

red green	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

The table shows all the possible outcomes.

Probability Distribution Example: *X* is the Sum of Two Dice

red / green	1	2	3	4	5	6
1	**2**	**3**	**4**	**5**	**6**	**7**
2	**3**	**4**	**5**	**6**	**7**	**8**
3	**4**	**5**	**6**	**7**	**8**	**9**
4	**5**	**6**	**7**	**8**	**9**	**10**
5	**6**	**7**	**8**	**9**	**10**	**11**
6	**7**	**8**	**9**	**10**	**11**	**12**

X
2
3
4
5
6
7
8
9
10
11
12

If you look at the table, you can see that *X* can be any of the numbers from 2 to 12.

Probability Distribution Example: X is the Sum of Two Dice

red / green	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

X	f
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	

We will now define *f*, the frequencies associated with the possible values of *X*.

Probability Distribution Example: *X* is the Sum of Two Dice

green \ red	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

X	f
2	
3	
4	
5	4
6	
7	
8	
9	
10	
11	
12	

For example, there are four outcomes which make *X* equal to 5.

Probability Distribution Example: *X* is the Sum of Two Dice

red / green	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

X	f
2	1
3	2
4	3
5	4
6	5
7	6
8	5
9	4
10	3
11	2
12	1

Similarly, you can work out the frequencies for all the other values of *X*.

Probability Distribution Example: X is the Sum of Two Dice

red / green	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

X	f	p
2	1	
3	2	
4	3	
5	4	
6	5	
7	6	
8	5	
9	4	
10	3	
11	2	
12	1	

Finally we will derive the probability of obtaining each value of *X*.

Probability Distribution Example: *X* is the Sum of Two Dice

red / green	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

X	f	p
2	1	
3	2	
4	3	
5	4	
6	5	
7	6	
8	5	
9	4	
10	3	
11	2	
12	1	

If there is 1/6 probability of obtaining each number on the red die, and the same on the green die, each outcome in the table will occur with 1/36 probability.

Probability Distribution Example: *X* is the Sum of Two Dice

red / green	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

X	*f*	*p*
2	1	1/36
3	2	2/36
4	3	3/36
5	4	4/36
6	5	5/36
7	6	6/36
8	5	5/36
9	4	4/36
10	3	3/36
11	2	2/36
12	1	1/36

Hence to obtain the probabilities associated with the different values of *X*, we divide the frequencies by 36.

Probability Distribution Example: *X* is the Sum of Two Dice

The distribution is shown graphically. in this example it is symmetrical, highest for *X* equal to 7 and declining on either side.

Overview

- Random Variables
- Expected Value
- Pairs of Discrete Random Variables
  - Conditional Probability
  - Bayes Rule
- Continuous Random Variables

Expected Value

- Definition of $E(X)$, the expected value of X:

$$E(X) = x_1 p_1 + \ldots + x_n p_n = \sum_{i=1}^{n} x_i p_i$$

- The expected value of a random variable, also known as its population mean, is the weighted average of its possible values, the weights being the probabilities attached to the values

Expected Value Example

x_i	p_i	$x_i p_i$
x_1	p_1	$x_1 p_1$
x_2	p_2	$x_2 p_2$
x_3	p_3	$x_3 p_3$
x_4	p_4	$x_4 p_4$
x_5	p_5	$x_5 p_5$
x_6	p_6	$x_6 p_6$
x_7	p_7	$x_7 p_7$
x_8	p_8	$x_8 p_8$
x_9	p_9	$x_9 p_9$
x_{10}	p_{10}	$x_{10} p_{10}$
x_{11}	p_{11}	$x_{11} p_{11}$

$$\Sigma\, x_i p_i = E(X)$$

x_i	p_i	$x_i p_i$
2	1/36	2/36
3	2/36	6/36
4	3/36	12/36
5	4/36	20/36
6	5/36	30/36
7	6/36	42/36
8	5/36	40/36
9	4/36	36/36
10	3/36	30/36
11	2/36	22/36
12	1/36	12/36

252/36 = 7

Expected Value Properties

- Linear

$$E(X + Y) = E(X) + E(Y)$$

$$E(bX) = bE(X)$$

$$E(b) = b$$

$$\begin{aligned} Y &= b_1 + b_2 X \\ E(Y) &= E(b_1 + b_2 X) \\ &= E(b_1) + E(b_2 X) \\ &= b_1 + b_2 E(X) \end{aligned}$$

- Also denoted by μ

Variance

Var(X) = E[(X- μ)^2] = $\sum (x_i - \mu)^2 P(X = x_i)$

Var(X) = σ^2

Ex: Find the mean and variance of the PMF.

x	1	2	3	4	5	6
$p(x)$	.30	.25	.15	.05	.10	.15

$$E(X) = \sum_{x=1}^{6} xp(x) = 2.85$$

$$Var(X) = \sum_{x=1}^{6} (x - \mu)^2 p(x) = 3.2275$$

Variance

$\mathrm{Var}(X) = E[(X - \mu)^2] = E[X^2] - (E[X])^2$

Proof:

$$\begin{aligned}
\mathrm{Var}(X) &= E[(X - E[X])^2] \\
&= E[X^2 - 2X\,E[X] + E[X]^2] \\
&= E[X^2] - 2\,E[X]^2 + E[X]^2 \\
&= E[X^2] - E[X]^2
\end{aligned}$$

Overview

- Discrete Random Variables
- Expected Value
- Pairs of Discrete Random Variables
  - Conditional Probability
  - Bayes Rule
- Continuous Random Variables

Pairs of Discrete Random Variables

- Let X and Y be two discrete r.v.
- For each possible pair of values, we can define a joint probability $p_{ij}=Pr[X=x_i, Y=y_j]$
- We can also define a joint probability mass function P(x,y) which offers a complete characterization of the pair of r.v.

$$P_X(x) = \sum_{y \in Y} P(x, y)$$

$$P_Y(y) = \sum_{x \in X} P(x, y)$$

Marginal distributions

Statistical Independence

Two random variables *x* and *y* are said to be independent, if and only if

$$P(x,y)=P_x(x)\,P_y(y)$$

that is, when knowing the value of *x* does not give us additional information for the value of *y*.

Or, equivalently

$$E[f(x)g(y)] = E[f(x)]\,E[g(y)]$$

for any functions f(x) and g(y).

Conditional Probability

- When two r.v. are not independent, knowing one allows better estimate of the other (e.g. outside temperature, season)

$$\Pr[x = x_i \mid y = y_j] = \frac{\Pr[x = x_i , y = y_j]}{\Pr[y = y_j]}$$

- If independent $P(x|y)=P(x)$

Law of Total Probability

$$P_x(x) = \sum_{y \in Y} P(x, y)$$

$$P(x \mid y) = \frac{P(x, y)}{P(y)}$$

Bayes Rule

$$P(x \mid y) = \frac{P(x, y)}{P(y)} = \frac{P(y \mid x)P(x)}{\sum_{x \in X} P(x, y)}$$

$$\text{posterior} = \frac{\text{likelihood} * \text{prior}}{\text{evidence}}$$

- *x* is the unknown cause
- y is the observed evidence
- Bayes rule shows how probability of *x* changes after we have observed *y*

Overview

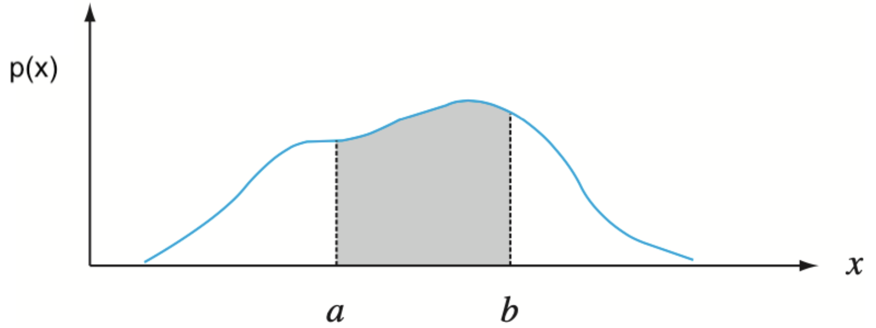
- Discrete Random Variables
- Expected Value
- Pairs of Discrete Random Variables
  - Conditional Probability
  - Bayes Rule
- Continuous Random Variables

Continuous Random Variables

- Examples: room temperature, time to run 100m, weight of child at birth…
- Cannot talk about probability of that X has a particular value
- Instead, probability that X falls in an interval => probability density function

$$\Pr[x \in (a,b)] = \int_a^b p(x)dx$$

$$p(x) \geq 0 \text{ and } \int_{-\infty}^{\infty} p(x)dx = 1$$



$P(a \leq X \leq b)$ = the area under the density curve between a and b

Probability and statistics for engineering and the sciences – Jay Devore

Cumulative Distribution Functions

The cumulative distribution function (CDF) of a random variable X is denoted by F_X and provides the probability $P(X \leq x)$.

$$F_X(x) = P(X \leq x) = \begin{cases} \sum_{k \leq x} p_X(k) & X: \ discrete \\ \int_{-\infty}^{x} f_X(t)dt & X: \ continuous \end{cases}$$

Expected Value

$$E[x] = \mu = \int_{-\infty}^{\infty} x p(x) dx$$

$$E[f(x)] = \int_{-\infty}^{\infty} f(x) p(x) dx$$

$$Var[x] = \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 p(x) dx$$

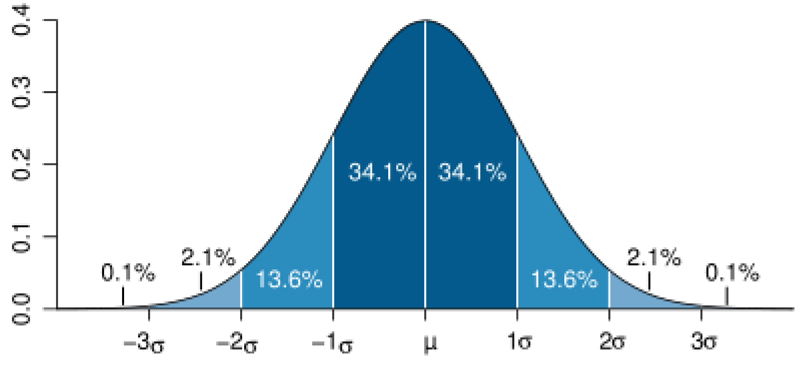
- Bayes rule $$p(x \mid y) = \frac{p(y \mid x) p(x)}{\int_{-\infty}^{\infty} p(y \mid x) p(x) dx}$$

$$\text{posterior} = \frac{\text{likelihood} * \text{prior}}{\text{evidence}}$$

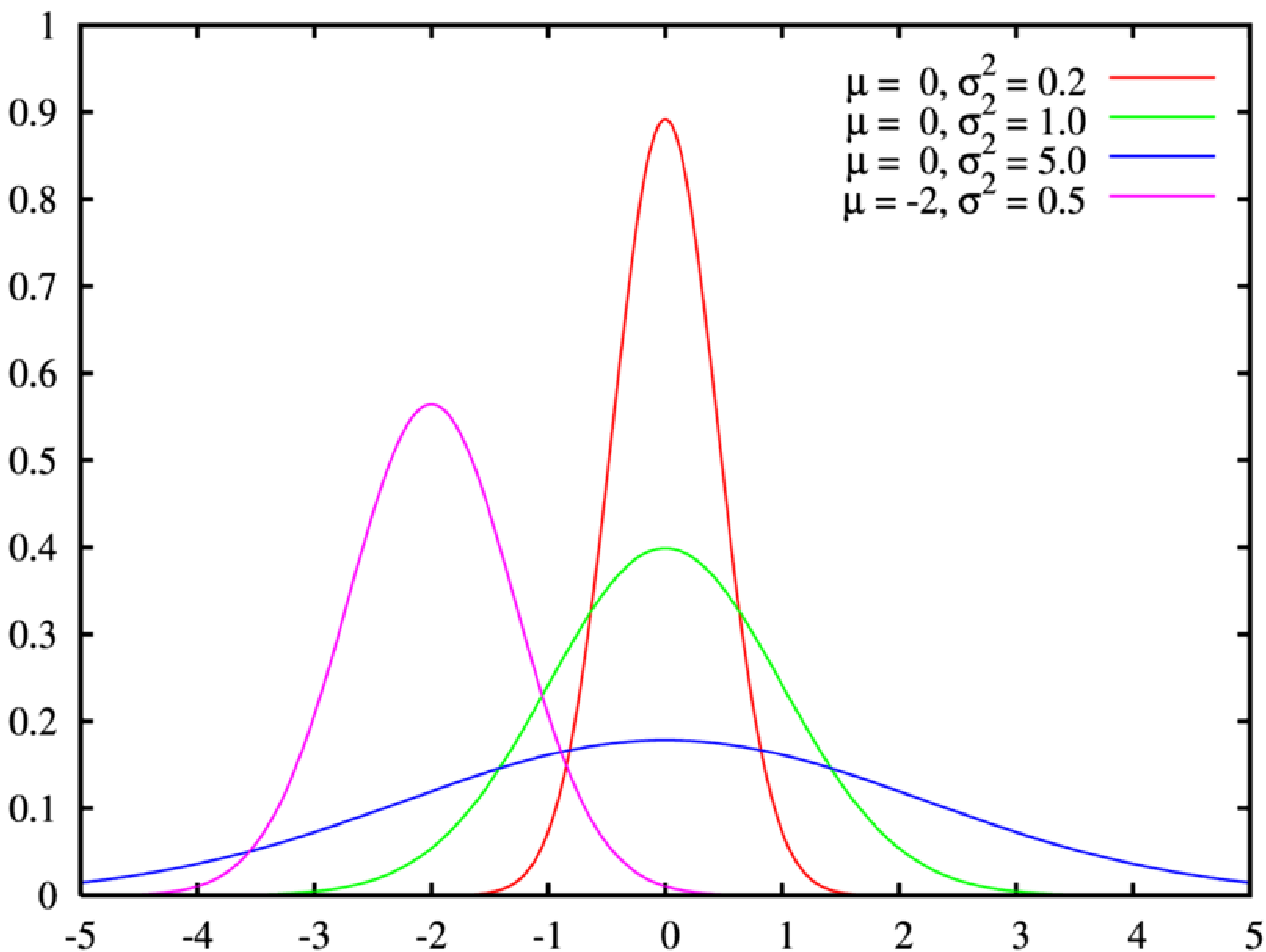
Normal Random Variables

- A continuous random variable X is said to be normal or Gaussian if it has a PDF of the form:

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



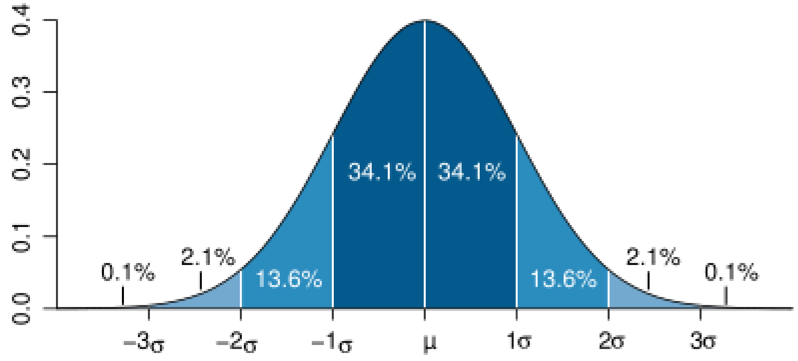
Normal (Gaussian) Distribution



Normal (Gaussian) Distribution

- Central Limit Theorem: under various conditions, the distribution of the sum of *d* independent random variables approaches a limiting form known as the normal distribution

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



Z-score

A z-score (also called a *standard score*) gives you an idea of how far from the mean a data point is. But more technically it's a measure of how many standard deviations below or above the population mean a raw score is.

A z-score can be placed on a normal distribution curve.

$$z = \frac{x - \mu}{\sigma}$$

https://www.statisticshowto.com/probability-and-statistics/z-score/

Standard Normal z-table

The z-table shows what percentage is under the curve at any particular point.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

https://www.ztable.net/

Example

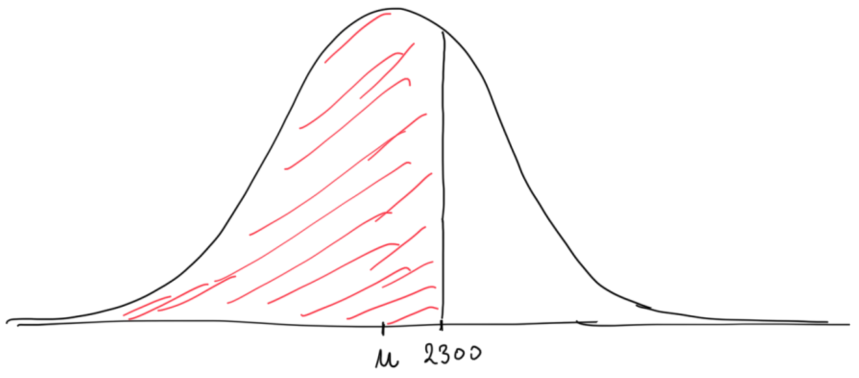
A coffee shop in Hoboken buys the coffee beans in Brazil. The owner of the shop likes to predict the price of the coffee beans one month in advance while she prepares budget plannings. For predicting the price of the beans, the owner uses the normal distribution for the model where the mean is 2250 dollars and standard deviation is 110 dollars. Help the little coffee shop in Hoboken to answer the following questions.

Example

What is the probability that the coffee beans market price one month from now wound not exceed 2300 dollars?

$$z = \frac{x - \mu}{\sigma} = \frac{2300 - 2250}{110} = 0.454545$$

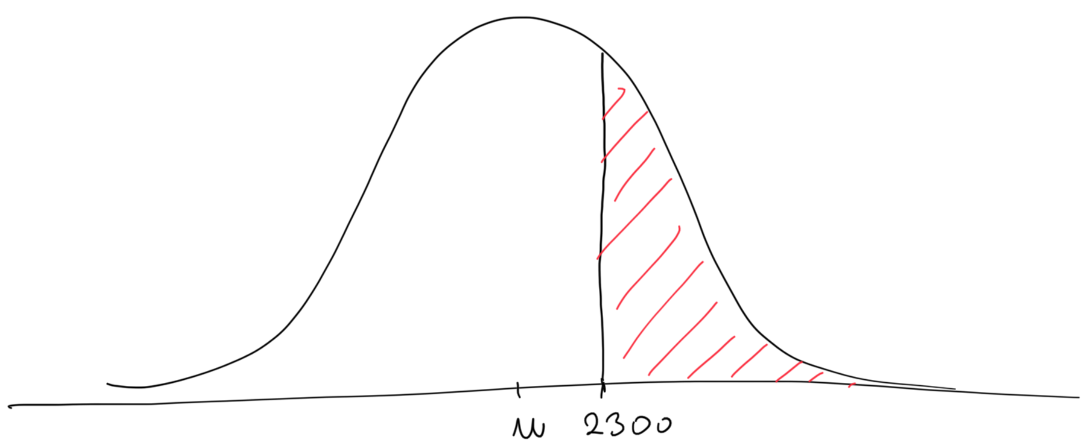
$$P(Price \leq 2300) = 0.6736$$



Example

What is the probability that the market price one month from now exceeds 2300 dollars?

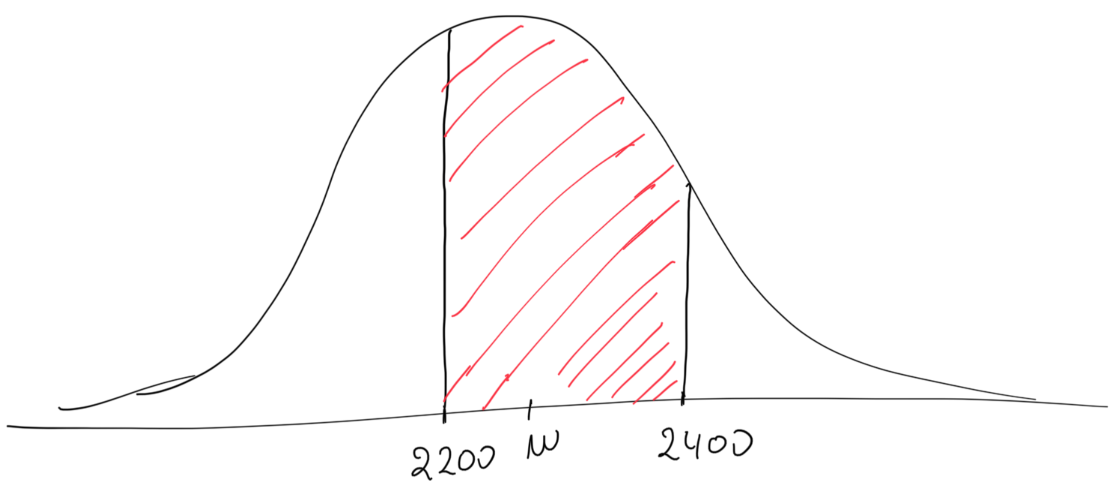
$$P(Price > 2300) = 1 - P(Price \leq 2300) = 1 - 0.6736 = 0.3264$$



Example

What is the probability that the market price one month from now will be between 2200 and 2400 dollars?

$$P(2200 < Price < 2400) = P(Price < 2400) - P(Price < 2200) = 0.5867$$



Thank you!