UNIT 2

Identify whether the following are Discrete or Continuous Random Variables

1. Number of free throws an NBA player makes in his next 20 attempts.

2. Time between lightening strikes in a thunderstorm.

3. Velocity of next pitch in Major League Baseball.

4. Number of rolls of a die needed to roll a 3 for first time.

Linear Transformation of Random Variables

SRS travels offers a half-day trip in a tourist area. There must be at least 2 passengers for the visit to run.
The vehicle provided by SRS travels can hold up to 6 passengers.
Let X represent the No. Of passengers that turn up on a randomly selected day.

Probability Distribution of X

X	2	3	4	5	6
p(x)	0.15	0.25	0.35	0.20	0.05

a. i) Calculate average number of passengers that turn up for the visit.
   ii) Calculate SD.

Mean = 3.75 standard dev = 1.090

b. SRS travels charges Rs. 150 per passenger.

Let Y represent the amount SRS travels collects on a randomly selected day.

Provide probability distribution of Y.

Calculate mean and SD of Y.

Y	300	450	600	750	900
p(x)	0.15	0.25	0.35	0.20	0.05

Mean = 3.75 standard dev = 1.090 (for X)
Mean = 562.50 standard dev = 163.50 (for Y)

c.The amount spent on petrol and permit by SRS travels per trip is Rs. 100.
Let Z represent the Profit made by SRS travels on a randomly selected day.
Find the probability distribution of Z. Find mean and SD of Z.

Z	200	350	500	650	800
p(x)	0.15	0.25	0.35	0.20	0.05

Mean = 3.75 standard dev = 1.090 (for X)
Mean = 462.50 standard dev = 163.50 (for Z)

Chebyshev's Inequality

Let X be a random variable with mean μ_X and standard deviation σ_X . Then

$$\mathbf{P(|X-\mu_X|\geq k\sigma_X)\leq 1/ k^2}$$

Examples

The length of a rivet manufactured by a certain process has mean $\mu_X = 50$ mm and standard deviation $\sigma_X = 0.45$ mm. What is the largest possible value for the probability that the length of the rivet is outside the interval 49.1–50.9 mm?

Solution
Let X denote the length of a randomly sampled rivet.
We must find $P(X \leq 49.1 \text{ or } X \geq 50.9)$.

Now
$P(X \leq 49.1 \text{ or } X \geq 50.9) = P(|X - 50| \geq 0.9) = P(|X - \mu_X | \geq 2\sigma_X)$

Applying Chebyshev's inequality with k = 2, we conclude that

$P(X \leq 49.1 \text{or } X \geq 50.9) \leq 1/4$

Example #2

A class of second graders has a mean height of five feet with a standard deviation of one inch. At least what percent of the class must be between 4’10” and 5’2”?

Solution

The heights that are given in the range above are within two standard deviations from the mean height of five feet. Chebyshev’s inequality says that at least $1 – 1/2^2 = 3/4 = 75\%$ of the class is in the given height range.

Example #3

Computers from a particular company are found to last on average for three years without any hardware malfunction, with a standard deviation of two months. At least what percent of the computers last between 31 months and 41 months?

Solution

The mean lifetime of three years corresponds to 36 months. The times of 31 months to 41 months are each $5/2 = 2.5$ standard deviations from the mean. By Chebyshev’s inequality, at least $1 – 1/(2.5)6^2 = 84\%$ of the computers last from 31 months to 41 months.

Example #3

Bacteria in a culture live for an average time of three hours with a standard deviation of 10 minutes. At least what fraction of the bacteria live between two and four hours?

Solution

Two and four hours are each one hour away from the mean. One hour corresponds to six standard deviations. So at least $1 – 1/6^2 = 35/36 =97\%$ of the bacteria live between two and four hours.

Example #4

What is the smallest number of standard deviations from the mean that we must go if we want to ensure that we have at least 50% of the data of a distribution?

Solution

Here we use Chebyshev’s inequality and work backward. We want $50\% = 0.50 = 1/2 = 1 – 1/K^2$. The goal is to use algebra to solve for K.

We see that $1/2 = 1/K^2$. Cross multiply and see that $2 =K^2$. We take the square root of both sides, and since K is a number of standard deviations, we ignore the negative solution to the equation. This shows that K is equal to the square root of two. So at least 50% of the data is within approximately 1.4 standard deviations from the mean.

Example #5

Bus route #25 takes a mean time of 50 minutes with a standard deviation of 2 minutes. A promotional poster for this bus system states that "95% of the time bus route #25 lasts from _____ to ______ minutes." What numbers would you fill in the blanks with?

Solution

This question is similar to the last one in that we need to solve for K, the number of standard deviations from the mean. Start by setting 95% = 0.95 = $1 - 1/K^2$. This shows that $1 - 0.95 = 1/K^2$. Simplify to see that $1/0.05 = 20 = K^2$. So $K = 4.47$.

Now express this in the terms above. At least 95% of all rides are 4.47 standard deviations from the mean time of 50 minutes. Multiply 4.47 by the standard deviation of 2 to end up with nine minutes. So 95% of the time, bus route #25 takes between 41 and 59 minutes.

Sampling Distribution

A process that fills plastic bottles with a beverage has a mean fill volume of 2.013 L and a standard deviation of 0.005 L. A case contains 24 bottles. Assuming that the bottles in a case are a simple random sample of bottles filled by this method, find the mean and standard deviation of the average volume per bottle in a case.

Solution

Let $V_1, \ldots, V_{24}$ represent the volumes in 24 bottles in a case.

This is a simple random sample from a population with mean $\mu = 2.013$ and standard deviation $\sigma = 0.005$.

The average volume is $V = (V_1 + \cdots + V_{24})/24$.
Using Equation (2.56),
Using Equation (2.58),

$\mu_V = \mu = 2.013$

$\sigma_V = \sigma / \sqrt{24} = 0.001$

Discrete Distributions

Bernoulli Distribution

Problem 1:

If the probability of a bulb being defective is 0.8, then what is the probability of the bulb not being defective.

Solution:

Probability of bulb being defective, p = 0.8

Probability of bulb not being defective, q = 1 - p = 1 - 0.8 = 0.2

2. Approximately 1 in 200 adults are lawyers. One adult is randomly selected.

What is the distribution of the number of lawyers?

X – represents an adult is a lawyer.
X ~ Bernoulli (1 /200)

Probability Distribution of X

X	p(x)
0	199/200
1	1/200

3. Candidate A is running for office in a certain district.

Twenty persons are selected at random from the population of registered voters and asked if they prefer candidate A.

Do the responses form a sequence of Bernoulli trials?
If so identify the trial outcomes and the meaning of the parameter p.

Yes, the responses form a sequence of Bernoulli trials.
The outcomes are: prefer A and do not prefer A;
p is the proportion of voters in the entire district who prefer A

Suppose that a student takes a multiple choice test.
The test has 10 questions, each of which has 4 possible answers(only one is correct).
If the student blindly guesses the answer to each question, do the questions form a sequence of Bernoulli trials? If so, identify the trial outcomes and the parameter p.

For each question,
Either the answer chosen is correct or incorrect
P(Answer is correct) = $1/4$
P(Answer is incorrect) = $3/4$
Hence there are only 2 possible outcomes for each question. Hence each question is a Bernoulli trial.
Since there are in total 10 questions, we have a sequence of Bernoulli trials.

Binomial Distribution

Binomial Distribution

If a total of n Bernoulli trials are conducted, and

- The trials are independent
- Each trial has the same success probability p
- X is the number of successes in the n trials

then X has the binomial distribution with parameters n and p, denoted

$X \sim Bin(n, p)$.

Binomial Probability Formula

The binomial probability formula is the result of getting exactly r successes in n number of trials.

The Binomial probability formula is given as follow:

P(r success in n trials) = $nCr\ p^r\ q^{n-r}$

Where,

p = the probability of success and

q = the probability of failure (or complement of the event)

n = Total number of trials

r = number of specific events we want to obtain

Also nCr represents selection of r events from n, it can be written as: $nCr = n! * r! / (n-r)!$

If a trial is done n times, then to **find the probability of success happening r times, probability distribution formula is used.**

The Bernoulli trial binomial distribution formula is given as:

$P(X=r)=nC_r\, p^r q^{\,n-r}$

This formula can give solutions to problems like the **probability of r success in n trials, probability of at least one success in n trials, probability of no success in n trials, probability of at least one failure in n trials and probability of at most r success in n trials.**

Examples:

Ex 1:

Suppose, we are flipping a coin 4 times and we want to know the probability of getting a head 3 out of 4 times.

Ex 2:

If a student is attempting five true or false questions, then find the probability of getting 3 correct answers.

Ex 3:

Suppose, we are flipping a coin 4 times and we want to know the probability of getting a head 3 out of 4 times.

The probability of getting a head, p = 0.5

The probability of getting a tail, q = 0.5

The probability of getting a head 3 out of 4 times will be = pppq = 0.5 × 0.5 × 0.5 × 0.5

But out of 4 times, which trials will give head. For that, we will use combination. The 3 times when head will come out of 4 times can be arranged in $4C_3$ ways.

Hence, probability of getting 3 heads out of 4 will be $4C_3 *(0.5)^3 *(0.5)^1$

Ex 4:

If a student is attempting five true or false questions, then find the probability of getting 3 correct answers.

Probability of success, p = 0.5

Probability of failure, q = 0.5

Probability of getting 3 correct answers $=5\ C_3(0.5)^3(0.5)^2 = 0.3125$.

Ex 5 :

In a baseball game, the probability that John will get on base safely when he comes to bat is $\frac{3}{5}$. What is the probability that he will get on base safely at least 3 out of 4 times he comes to bat?

n = 4 because there are 4 trials

r = 3 and 4 because at least 3 means 3 or more

$$p = \frac{3}{5}$$

because it represents the probability of getting on base safely

$$q = 1 - \frac{3}{5} = \frac{2}{5}$$

because it represents the probability of not getting on base safely

P(getting on base safely in at least 3 out of 4 at bats) = $_4C_3\left(\frac{3}{5}\right)^3\left(\frac{2}{5}\right)^1 + {}_4C_4\left(\frac{3}{5}\right)^4\left(\frac{2}{5}\right)^0$

$$= \frac{216}{625} + \frac{81}{625} = \frac{297}{625}$$

Ex 6:

10 coins are tossed simultaneously where the probability of getting head for each coin is 0.6. Find the probability of getting 4 heads.

Solution:

Probability of getting head, p = 0.6

Probability of getting head, q = 1 - p = 1 - 0.6 = 0.4

Probability of getting 4 heads out of 10,

P(X=4)=10C4=(0.6)4(0.4)6

= 0.111476736

7. In an exam, 10 multiple choice questions are asked where only one out of four answers is correct. Find the probability of getting 5 out of 10 questions correct in an answer sheet.

Solution:

Probability of getting an answer correct, p = 1/4 = 0.25

Probability of getting an answer correct, q = 1 - p = 1 - 0.25 = 0.75

Probability of getting 5 answers correct, P(X = 5) =10 C5$(0.25)^5(0.75)^5$ = 0.05839920044

Problem 8:

The chances of a team winning a match is 0.7.

Find the probability that the team will win at least one match out of three.

Solution:

Probability of winning a match, p = 0.7

Probability of winning a match, q = 1 - p = 1 - 0.7 = 0.3

Probability of winning zero matches, P(X = 0) = C30(0.7)0(0.3)3 = 0.027

Probability of winning at least one match = 1 - 0.027 = 0.973

Using a Sample Proportion to Estimate a Success Probability

In many cases we do not know the success probability p associated with a certain Bernoulli trial, and we wish to estimate its value. A natural way to do this is to conduct n independent trials and count the number X of successes. To estimate the success probability p we compute the sample proportion p.

p = number of successes / number of trials =X / n

This notation follows a pattern that is important to know. The success probability, which is unknown, is denoted by p. The sample proportion, which is known, is denoted p. The "hat" ( ) indicates that p is used to estimate the unknown value p.

A quality engineer is testing the calibration of a machine that packs ice cream into containers. In a sample of 20 containers, 3 are underfilled. Estimate the probability p that the machine underfills a container.

Solution

The sample proportion of underfilled containers is p = 3/20 = 0.15. We estimate that the probability p that the machine underfills a container is 0.15 as well.

Uncertainty in the Sample Proportion

It is important to realize that the sample proportion *p hat* is just an *estimate* of the success probability *p*, and in general, is *not equal* to *p*. If another sample were taken, the value of *p hat* would probably come out differently. In other words, there is uncertainty in *p hat* .

For *p* to be useful, we must compute its bias and its uncertainty.

We now do this.

Let *n* denote the sample size, and let *X* denote the number of successes, where $X \sim \text{Bin}(n, p)$.

The bias is the difference $\mu_{p\ hat} - p$. Since $p = X/n$, it follows that

$\mu_{p\ hat} = \mu_{X/n} = 1/n * \mu_X = 1/n * np = p$

Since $\mu_{p\ hat} = p$, p is unbiased; in other words, its bias is 0.

The uncertainty is the standard deviation $\sigma_{p\ hat}$.

The standard deviation of X is $\sigma_X = \sqrt{np(1 - p)}$. Since $p\ hat = X/n$, it follows that

$\sigma_{p\ hat} = \sigma_{X/n} = 1/n * \sigma_X = 1/n * \sqrt{np(1- p)} = \sqrt{p(1- p)/n}$

In practice, when computing the uncertainty in p, we don't know the success probability p, so we approximate it with $p\ hat$.

The Poisson Probability Distribution

- Mean and variance of Poisson distribution

The Poisson Distribution was developed by the French mathematician Simeon Denis Poisson in 1837.

Poisson was a French mathematician, and amongst the many contributions he made, proposed the Poisson distribution, with the example of modeling the number of soldiers accidentally injured or killed from kicks by horses. This distribution became useful as it models events, particularly uncommon events.

Poisson Distribution – Examples

1) The number of deaths by horse kicking in the Prussian army (First application).
2) The number of cyclones in a season.
3) Arrival of Telephone calls, Customers, Traffic, Web requests.
4) Estimating the number of mutations of DNA after exposure to radiation.
5) Rare diseases (like Leukemia(cancer of the blood cells), but not AIDS because it is infectious and so not independent).
6) The number of calls coming per minute into a hotels reservation Center.
7) The number of particles emitted by a radioactive source in a given time.
8) The number of births per hour during a given day.
9) The number of patients arriving in an emergency room between 11 and 12 pm.
10) The number of car accidents in a day.

11)birth defects and genetic mutations

12)number of typing errors on a page

13)hairs found in McDonald's hamburgers

14)spread of an endangered animal in India

15)failure of a machine in one month

The Poisson random variable satisfies the following conditions:

1. The number of successes in two disjoint time intervals is independent.

2. The probability of a success during a small time interval is proportional to the entire length of the time interval.

The **probability distribution of a Poisson random**

variable *X* representing the number of successes occurring in a given time interval or a specified region of space is given by the formula:

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

Where x=0,1,2,3… and e=2.71828 (but use your calculator's e button)

μ= mean number of successes in the given time interval or region of space

Properties of Poisson distribution

Events occur independently

The probability that an event occurs in a given interval of time is constant (does not change with time)

Events occur randomly andindependently.

Then, X, the number of events in a fixed unit of time, has a Poisson distribution.

$$P(X=x)= e^{-\lambda} \lambda^{x} \; / \; x!$$

Where λ is mean of the distribution or the average rate at which events are occuring in a given interval of time or space.

Mean and Variance of Poisson distribution

If μ is the average number of successes occurring in a given time interval or region in the Poisson distribution, then the mean and the variance of the Poisson distribution are both equal to μ.

$E(X) = \lambda$

and

$V(X) = \sigma^2 = \lambda$

Note: In a Poisson distribution, only one parameter, μ is needed to determine the probability of an event.

Example 1

If electricity power failures occur according to a Poisson distribution with an average of 3 failures every twenty weeks, calculate the probability that there will not be more than one failure during a particular week.

Average no of failures per 20 weeks = λ = 3
X denote the number of failures per week
X ~ Poisson(λt)
X ~ Poisson(3/ 20) => X ~ Poisson(0.15)

2. Vehicles pass through a junction on a busy road at an average rate of 300 per hour.

1)Find the probability that none passes in a given minute.
2)What is the expected number passing in two minutes?
3)Find the probability that this expected number actually pass through in a given two-minute period.

Average no of vehicles passing per hour =300
Let X denote Average no of vehicles passing per minute
X ~ Poisson(λt) => X ~ Poisson(300 / 60) => X ~ Poisson(5)

1) P(none passes in a given minute) = P(X = 0) =6,7379 X 10^{-3}
2) Expected no passing in two minutes = E(X) = 5 * 2 = 10.
3) Average no of vehicles that pass in 2 min period = λ = 10
4) P(10 passes in a given minute) = P(X = 10) =0.12511

A life insurance salesman sells on the average 3 life insurance policies per week.

Find the probability that in a given week he will sell

a. Some policies
b. 2 or more policies but less than 5 policies.
c. Assuming that there are 5 working days per week, what is the probability that in a given day he will sell one policy?

λ =3
1) Some policies means P(X > 0)
P(X > 0) = 1 – P(X = 0) = 1- 0.04 = 0.96

2) 2 or more policies but less than 5 policies.
P(2 <= X < 5) = P(X =2) + P(X =3) + P(X = 4)
=0.61611

3) Assuming that there are 5 working days per week, what is the probability that in a given day he will sell one policy?

Avg. no. policies sold per day = 3/5 =0.6

$P(X = 1) = 0.32929$

Grandma bakes chocolate chip cookies in batches of 100. She puts 300 chips into the dough. When the cookies are done, she gives you one. What is the probability that your cookie contains no chocolate chips?

Solution

Let X represent the number of chips in your cookie.
The mean number of chips is 3 per cookie, so $X \sim \text{Poisson}(3)$.
It follows that $P(X = 0) = e^{-3}3^0/0! = 0.0498$.

Grandma's grandchildren have been complaining that Grandma is too stingy with the chocolate chips. Grandma agrees to add enough chips to the dough so that only 1% of the cookies will contain no chips. How many chips must she include in a batch of 100 cookies to achieve this?

Solution

Let n be the number of chips to include in a batch of 100 cookies, and

let X be the number of chips in your cookie.

The mean number of chips is 0.01n per cookie, so
$X \sim \text{Poisson}(0.01n)$.

We must find the value of n for which $P(X = 0) = 0.01$.

Using the Poisson(0.01n) probability mass function,
$P(X = 0) = e^{-0.01n} (0.01n)^0 / 0!$

we obtain $n \approx 461$.

Examples

A man was able to complete 3 files a day on an average. Find the probability that he can complete 5 files the next day.

Solution: Here we know this is a Poisson experiment with following values given:

$\mu = 3$, average number of files completed a day

$x = 5$, the number of files required to be completed next day

And $e = 2.71828$ being a constant

On substituting the values in the Poisson distribution formula mentioned above we get the Poisson probability in this case.

$P(x, \mu) = ((e^{-\mu})(\mu^{x}) / x!$

$P(5, 3) = (2.71828)^{-3}(3^{5}) / 5!$

$= 0.1008$ approximately.

Hence the probability for the person to complete 5 files the next day is 0.1008 approximately.

Practice Problems

Patients arrive at a hospital accident and emergency department at random at a rate of 6 per hour.

(a) Find the probability that, during any 90 minute period, the number of patients arriving at the hospital accident and emergency department is

(i) exactly 7

(ii) at least 10

(5)

A patient arrives at 11.30 a.m.

(b) Find the probability that the next patient arrives before 11.45 a.m.

(3)

An online shop sells a computer game at an average rate of 1 per day.

(a) Find the probability that the shop sells more than 10 games in a 7 day period.

(3)

Once every 7 days the shop has games delivered before it opens.

(b) Find the least number of games the shop should have in stock immediately after a delivery so that the probability of running out of the game before the next delivery is less than 0.05

(3)

In an attempt to increase sales of the computer game, the price is reduced for six months. A random sample of 28 days is taken from these six months. In the sample of 28 days, 36 computer games are sold.

(c) Using a suitable approximation and a 5% level of significance, test whether or not the average rate of sales per day has increased during these six months. State your hypotheses clearly.

(7)

In a village, power cuts occur randomly at a rate of 3 per year.

(a) Find the probability that in any given year there will be

(i) exactly 7 power cuts,

(ii) at least 4 power cuts.

(5)

(b) Use a suitable approximation to find the probability that in the next 10 years the number of power cuts will be less than 20

(6)

The number of houses sold by an estate agent follows a Poisson distribution, with a mean of 2 per week.

(a) Find the probability that in the next 4 weeks the estate agent sells,

(i) exactly 3 houses,

(ii) more than 5 houses.

(5)

The estate agent monitors sales in periods of 4 weeks.

(b) Find the probability that in the next twelve of these 4 week periods there are exactly nine periods in which more than 5 houses are sold.

(3)

The estate agent will receive a bonus if he sells more than 25 houses in the next 10 weeks.

(c) Use a suitable approximation to estimate the probability that the estate agent receives a bonus.

(6)

Defects occur at random in planks of wood with a constant rate of 0.5 per 10 cm length. Jim buys a plank of length 100 cm.

(a) Find the probability that Jim's plank contains at most 3 defects.
(2)

Shivani buys 6 planks each of length 100 cm.

(b) Find the probability that fewer than 2 of Shivani's planks contain at most 3 defects.
(5)

(c) Using a suitable approximation, estimate the probability that the total number of defects on Shivani's 6 planks is less than 18.
(6)

Normal Distribution

Examples

The annual salaries of employees in a large company are approximately normally distributed with a mean of $50,000 and a standard deviation of $20,000.

a) What percent of people earn less than $40,000?

b) What percent of people earn between $45,000 and $65,000?

c) What percent of people earn more than $70,000?

a) For x = 40000, z = -0.5

Area to the left (less than) of z = -0.5 is equal to 0.3085 = 30.85% earn less than $40,000.

b) For x = 45000, z = -0.25 and for x = 65000, z = 0.75

Area between z = -0.25 and z = 0.75 is equal to 0.3720 = 37.20 earn between $45,000 and $65,000.

c) For x = 70000, z = 1

Area to the right (higher) of z = 1 is equal to 0.1586 = 15.86% earn more than $70,000.

Ex 2:

The lifetime of a battery is in a certain application is normally distributed with mean 16 hours, standard deviation 2 hours.

a) What is the probability that a battery will last more than 19 hours?

b) Find the 10th percentile of the lifetimes.

c) A particular battery lasts 14.5 hours. What percentile is its lifetime on?

The lifetime of a battery is in a certain application is normally distributed with mean 16 hours, standard deviation 2 hours.

a) What is the probability that a battery will last more than 19 hours?

To find: $P(X \geq 19)$

$Z = (19 – 16)/2 = 1.5$
$P(Z \geq 1.5)=1–P(Z\leq 1.5)=1–0.9332=0.0668$

b) Find the 10th percentile of the lifetimes.

$P(Z \leq c) = 0.1000$
Closest area to 0.1000 is 0.1003 corresponding to a z-score of -1.28 =>
$-1.28 = (x – 16)/2$ => $x = 13.44$

c) A particular battery lasts 14.5 hours.
What percentile is its lifetime on?

$Z = (14.5 – 16) / 2 = -0.75$
$P(Z \leq -0.75) = 0.2266$
=> its lifetime is approximately on 23rd percentile.

Ex 3:

Scores on a standardized test are approximately normally distributed with mean 480 and a standard deviation of 90.

a) What proportion of the scores are above 700?
b) What is the 25th percentile of the scores?

c) If Arun's score is 600, what percentile is he on?

d) What proportion of the scores are between 420 and 520?

a) What proportion of the scores are above 700?
$z = (700 - 480)/90 = 2.44$.
The area to the right of $z = 2.44$ is 0.0073.

b) What is the 25th percentile of the scores?
The z-score of the 25th percentile is ≈ -0.67.
The 25th percentile is therefore $\approx 480 - 0.67(90) = 419.7$

c) If Arun's score is 600, what percentile is he on?

$z = (600 - 480)/90 = 1.33$.
The area to the left of $z = 1.33$ is 0.9082. Therefore a score of 600 is on the 91st percentile, approximately.

d) What proportion of the scores are between 420 and 520?

For 420, $z = (420 - 480)/90 = -0.67$.

For 520, $z = (520 - 480)/90 = 0.44$.

The area between $z = -0.67$ and $z = 0.44$ is $0.6700 - 0.2514 = 0.4186$.

Sum/ Difference of two independent normally distributed random variables is normal

If $X \sim N(\mu, \sigma^2)$ and $Y \sim N(\mu, \sigma^2)$ are independent random variables that are normally distributed, then their sum/difference is also normally distributed.

i.e., If , $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$

Then,

$$X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$$

Similarly, $$X - Y \sim N(\mu_X - \mu_Y, \sigma_X^2 + \sigma_Y^2)$$

Linear Function of a Normal Random Variable

If $X \sim N(\mu, \sigma^2)$ and a and b are constants, and $Y = aX + b$ then,

$Y \sim N(a\mu_X + b, a^2\sigma_X^2)$

Example:

Let X1 be a normal random variable with mean 2 and variance 3,
and
let X2 be a normal random variable with mean 1 and variance 4.
Assume that X1 and X2 are independent.
What is the distribution of the linear combination Y = 2X1 + 3X2?

Y is normally distributed with mean 7 and variance 48 as the following calculation illustrates:
(2X1 + 3X2)~N(2(2)+3(1), 22(3)+32(4))= N(7,48)

Ex. 2:

A light fixture holds two light bulbs.

Bulb A is a type whose lifetime is normally distributed with mean 800 hours and standard deviation 100 hours.
Bulb B has a lifetime that is normally distributed with mean 900 hours and standard deviation 150 hours.
Assume the lifetimes of the bulbs are independent.

1) What is the probability Bulb B lasts longer than bulb A?

2) What is the probability Bulb B lasts 200 hours more than bulb A?

3) Another light fixture holds only one bulb. A bulb of type A is installed, and when it burns out, a bulb of type B is installed.
What is the probability that the total lifetime of the two bulbs is more than 2000 hours?

Bulb A : X ~ N(800, 1002)
Bulb B : Y ~ N(900, 1502)

Assume the lifetimes of the bulbs are independent.

1) What is the probability Bulb B lasts longer than bulb A?

D = Y – X.
The event B > A is the event D > 0
D ~ N(μX - μY , σX 2 + σY 2)
D ~ N (100, 180.282)

Since D is a linear combination of independent normal random variables,
D is normally distributed.

P(D > 0)
Z = (0 – 100)/180.28 => z = -0.55
P(Z > -0.55) = 1 – P(Z < -0.55)

= 1 – 0.2912 = 0.7088

2) What is the probability Bulb B lasts 200 hours more than bulb A?

P(D > 200)
Z = (200 – 100)/180.28 => z = 0.55
P(Z > 0.55) = 1 – P(Z < 0.55) = 1 – 0.7088 = 0.2912

3) Another light fixture holds only one bulb. A bulb of type A is installed, and when it burns out, a bulb of type B is installed.

What is the probability that the total lifetime of the two bulbs is more than 2000 hours?

$X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$
$X + Y \sim N(1700, 180.282)$

Since X + Y is a linear combination of independent normal random variables,
X + Y is normally distributed.

$P(X+Y>2000)$
$Z = (2000 - 1700)/180.28 \Rightarrow z = 1.66$
$P(Z > 1.66) = 1 - P(Z < 1.66) = 1 - 0.9515 = 0.0485$