STATS 260 Class 14

Gavin Jaeger-Freeborn

1. Sets 15 and 16

2. Normal Density Function

if X is normally distributed with the mean μ and standard deviation σ. then we wright $X \sim N(\mu, \sigma)$. The pdf of X is:

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

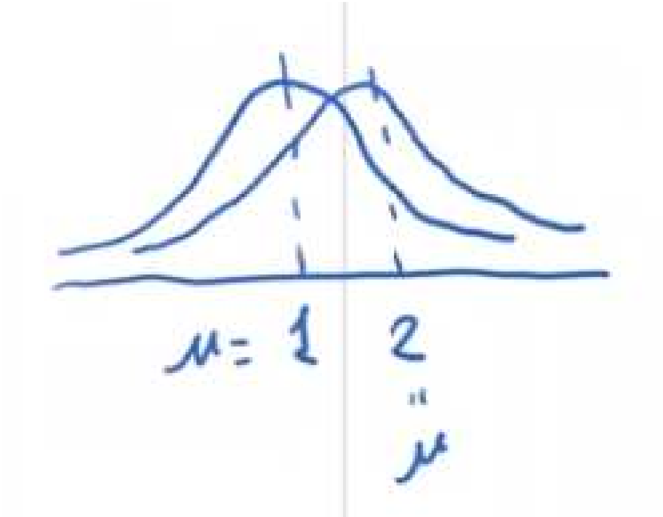
Properties of Normal Curves:

- All normal curves are defined on $(-\infty, \infty)$ and is bell-shaped.
- There is a single peak at $\underline{x = \mu}$ and the curve is symmetric about this peak.

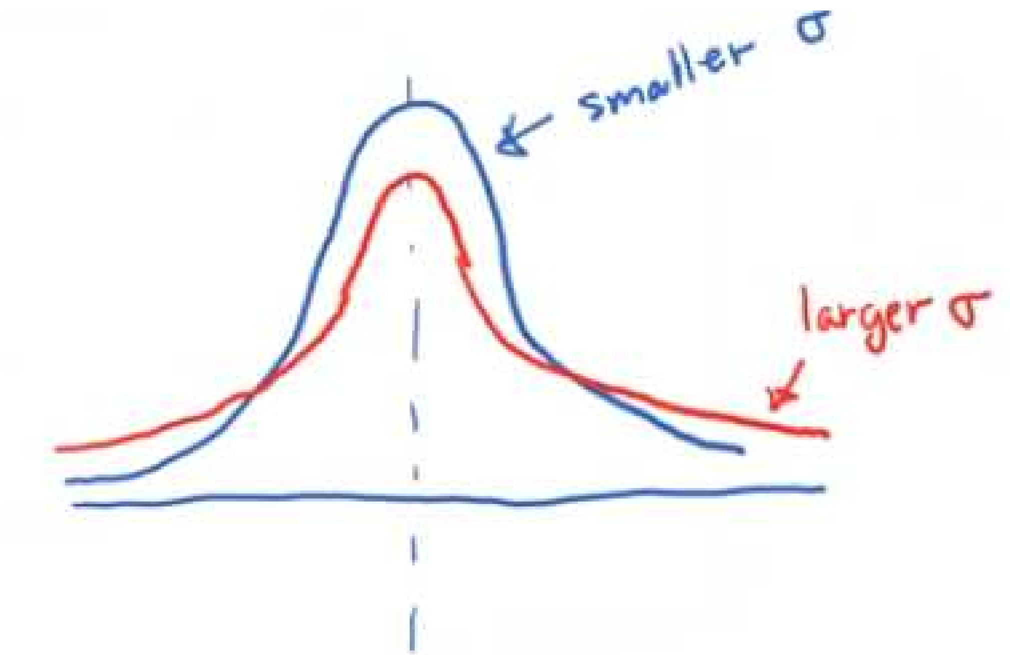
  μ is the max of $f(x)$

- The mean, median, and mode are all μ; the variance is σ 2 .
- There are points of inflection at $\mu - \sigma$ and $\mu + \sigma$

- As μ increases, the peak moves further to the right. As μ decreases, the peak moves further to the left. (μ is a **location parameter**)



- As σ increases, the peak becomes lower, and the curve becomes flatter. As σ decreases, the curve becomes more abruptly peaked, and the peak becomes taller. (σ is a **scale** parameter).



$$E(X) = \mu$$

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

3. Standard Normal Distribution

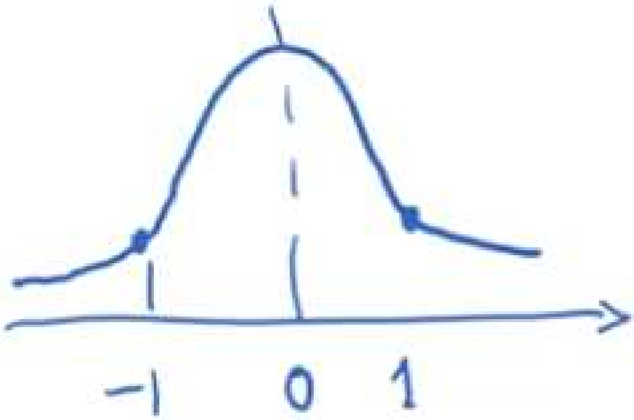
the standard normal random variable has mean 0 and standard deviation 1. We use the letter Z to denote the standard normal distribution.

$$N(0,1),\ \mu = 0\ ,\ \sigma\ = 1$$

$$f(z;0,1) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}}$$

The standard normal curve is:

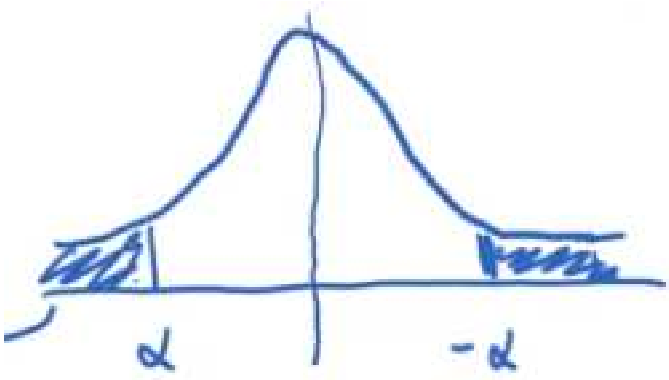
- has its peak at 0, and is symmetric about the y-axis
- has points of inflection at 1 and −1. If our random variable is Z, then we denote the cdf P $(Z \le z)$ by $\Phi(z)$.



4. Symmetry Property

Since the random variable Z is symmetric about $Z = 0$, then for any α

$$P(Z \leq \alpha) = P(Z \geq -\alpha)$$



Example

Find $P(Z \leq 2.56)$

Solution

$$\int_{-\infty}^{2.56} f(x)dx$$

USE THE TABLE

2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997
							0.06	

Table D.3 (continued) Areas under the Normal Curve

$$P(Z \leq 2.56) = 0.9948$$

or in R

```
> pnorm(2.56)
```

Example

Calculate $P(Z \geq 0.16)$

Solution

$$1 - P(Z \leq 0.16) \text{ or } P(Z \leq -0.16)$$

```
> 1 - pnorm(0.16)
[1] 0.4364405
```

$$\therefore P(Z \geq 0.16) = 0.4364405$$

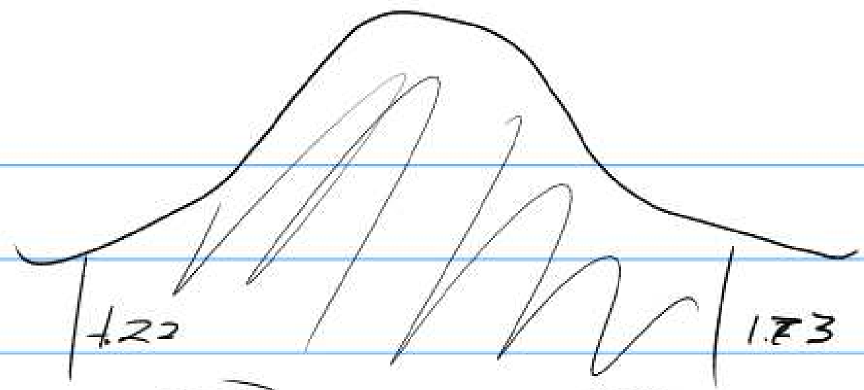
Example

Calculate $P(-1.22 < Z \leq 1.73)$

Solution

$$P(Z \leq 1.73) - P(Z \leq 1.22)$$

$$pnorm(1.73) - pnorm(-1.22)$$



$P(Z \leq -1.22)$ $P(Z \geq -1.22)$ $P(Z \leq 1.75)$

$P(Z \leq 1.75) - P(Z \leq -1.22)$

Example

Suppose that the heights of Andean flamingos are normally distributed with a mean of 105 cm and a standard deviation of 2 cm. Let the random variable X denote the height of a randomly selected Andean flamingo

What is the **median** Andean flamingo height?

$$\mu = 105, \ \sigma = 2$$

$$X \sim N(\mu = 105, \sigma = 2)$$

$$\boxed{105 \text{ cm}}$$

is $P(X \geq 100) = P(X \leq -100)$?

Only true for $Z \sim N(0, 1)$

$\therefore$ FALSE

What is $P(X = 105)$

As a property of all **continuous random variables**

$$P(X = 105) = 0$$

is $P(X \leq 100) = P(x \geq 110)$

$$\boxed{\text{True}}$$

Due to symmetry about its mean

Remember

$$P(X \leq \mu - x) = P(X \geq \mu + x)$$

NOTATION: z_α is the number such that $P(Z > z_\alpha) = \alpha$. Alternatively, z_α is the $100(1-\alpha)$ percentile of the standard normal distribution.

Example

Find the 97.5^{th} percentile of the standard normal distribution

$$100 - 97.5 = 2.5/100 = 0.025$$

$$Z_{0.025}$$

$$P(Z \leq Z_{0.025}) = 0.975$$

USE THE TABLE and find where the probability = 0.9750 and solve for z

$$Z_{0.025} = 1.96$$

in R

```
qnorm( 0.975 )
```

5. Standardizing a Normal Random Variable

If X is normally distributed with mean μ and standard deviation σ, then:

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

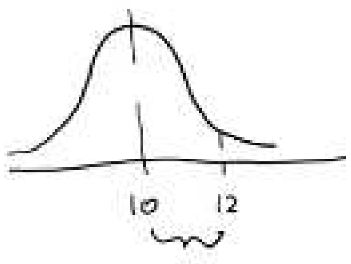
NOTE : this can be used for $P(Z \leq z)$ eg $X \sim N(105, 2)$

This basically means how many standard deviations away from the mean.

$$X = 12,\ \mu = 10\ \sigma = 2$$

$$Z = \frac{12 - 10}{2}$$

= 1 standard deviation from the mean



Example

The masses of a certain type of bolt is approximately normally distributed with μ = 15 g, and σ = 2 g. What is the probability that a **randomly selected bolt** has a mass between 14.3 g and 17.1 g?

$$P(14.3 \le X \le 17.1) = P(X \le 17.1) - P(X \le 14.3)$$

$$P(\frac{X-\mu}{\sigma} \le \frac{14.3-15}{2})$$

$$P(Z \le \frac{14.3-15}{2})$$

$$P(Z \le -0.35)$$

$$P(Z \le \frac{17.1-15}{2})$$

$$P(Z \le 1.05) - P(Z \le -0.35)$$

USING THE TABLE

$$= 0.8531 - 0.3632 = 0.4899$$

What is the probability that a randomly sleected bolt will have a mass of at least 20 g?

$$P(X \ge 20) = P(Z \ge \frac{20-15}{2})$$

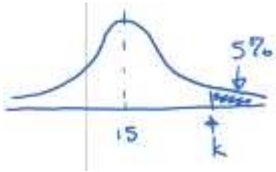
$$P(X \ge 20) = 1 - P(Z \le \frac{20-15}{2})$$

$$P(X \ge 20) = 1 - P(Z \le 2.5)$$

FROM TABLE

$$1 - 0.9938 = 0.0062$$

What is the minimum mass of the heavy 5% of bolts ?



$$P(X \leq k) = 9.5$$

$$P(Z \leq \frac{k - \mu}{\sigma}) = 9.5$$

REVERSE ON TABLE

$$P(Z \leq 1.645) = 0.95$$

$$\frac{k - 15}{2} = 1.645$$

solve for k

k = 2 (1.645) + 15 = 18.29

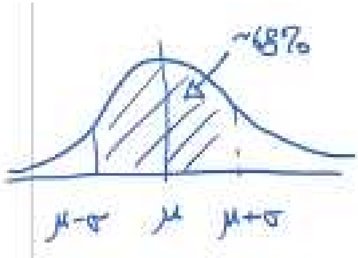
6. empirical rule

The empirical rule states that if the distribution of a variable is approximately normal, then:

1.
   About 68% of values lie within σ of the mean.

2.
   About 95% of values lie within 2σ of the mean.

1.
   About 99.7% of values lie within 3σ of the mean.



From this, we can conclude that almost all bolts will have a mass within 6g of the mean 15 (ie about 99.7% will have a mass between 9 g and 21 g).

$$\sigma = 2$$

9 ⟷ 15 ⟶ 21

$\frac{6}{2}$ $\frac{6}{2}$

3σ 3σ

(6 , 21) ≈ 99.7%

7. Approximating the Binomial Distribution with the Normal Distribution

Suppose $X \sim Bin(n, p)$ where np and $n(1-p)$ are both at least 5.

Then $X = N(\mu = np, \sigma^2 = np(1-p))$

This means that:

if

$np \geq 5$, and $n(1-p) \geq 5$

$$P(X \leq x) = P\left(Z \leq \frac{x - np}{\sqrt{np(1-p)}}\right)$$

Since we are using continuous distribution to approximate a discreet one , this approximation will be slightly off. If we with to get a better approximation us the following, with a **continuity correction**

$$P(X \leq x) = P(X \leq x + 0.5)$$

The $+0.5$ is for correction

$$= P\left(Z \leq \frac{x - np + 5}{\sqrt{np(1-p)}}\right)$$

eg

$$P(X \leq 3) = P(x \leq 3 + 0.5)$$

Example

Suppose it is known that 20% of batteries have a lifespan shorter than the advertised lifespan. Suppose that 100 batteries are selected at random.

What is the approximate probability (using the continuity correction) that at least 10 batteries will have a short lifespan?

$$X \sim Bin(100, 0.2)$$

$$np = 100(0.2) = 20$$

$$n(1-p) = 100(0.8) = 80$$

Both are ≥ 5

$$P(X >= 100)$$

$$= 1 - P(X \leq 9)$$

$$= 1 - P(X \leq 9 + 0.5)$$

$$X \sim N(20, \sqrt{100(0.2(0.8)})$$

$$= N(20, 4)$$

$$= 1 - P(X \leq 9.5)$$

$$= 1 - P(Z \leq -2.625)$$

USING THE TABLE

$$1 - 0.0043 = 0.9957$$

Example

Suppose it is known that the reaction time of type of voiceactivated robot is normally distributed with $\mu = 6.3$ microseconds, and $\sigma = 2$ microseconds.

Suppose I select one voice-activated robot at random. What is the probability that its reaction time is between 5 and 7 microseconds? Report your answer to three decimal places.

$$X \sim N(6.3, 2)$$

$$P(5 < X < 7)$$

$$= P(X < 7) - P(X < 5)$$

$$= P(Z < 0.35) - P(Z < -0.65)$$

FROM TABLE

$$= 0.6368 - 0.2578 = 0.379$$

Example

Suppose that I select five robots and test each of them. Assume the reaction time of each robot is independentof exactly three of the robots will have a reaction time between 5 and 7 microseconds? Report your answer to three decimal places.

Y = # of robots having reaction time between 5 and 7 microseconds

$$y \sim Bin(n = 5, p = 0.379)$$

$$P(y = 3) = \binom{5}{3} 0.379^3 (1 - 0.379)^2$$

$$= 0.210$$