ENM 207

Lecture 12

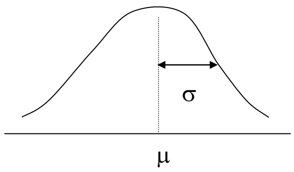
Some Useful Continuous Distributions

Normal Distribution

The most important continuous probability distribution in entire field of statistics.

Its graph, called the normal curve, is bell-shaped curve given below which describes approximately many phenomena occurring in

nature,
industry,
research



Normal Distribution

A continuous random variable *X* having bell-shape distribution is called a normal random variable.

The mathematical equation for the probability distribution of the normal random Variable depends upon two parameters μ and σ, its mean and standard deviation.

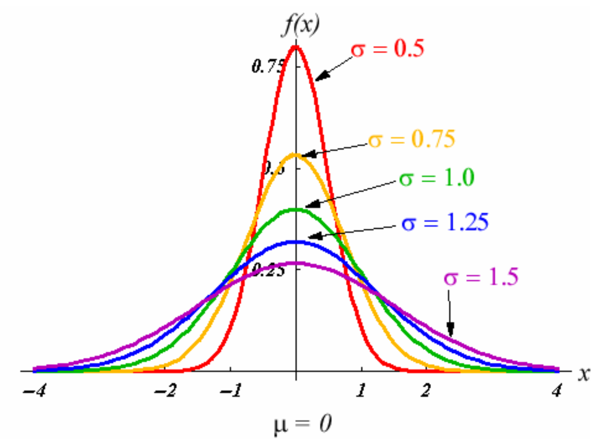
The density function of the normal random variable X, with mean μ and Variance σ^2, is:

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-(1/2)[(x-\mu)/\sigma]^2} & -\infty < x < \infty \\ 0 & \text{o.w.} \end{cases}$$

where π = 3.14159... and e = 2.71828…

Normal Distribution

A typical normal curves with different sigma (standard deviation) values are shown below;



Normal Distribution

Many population and process variables have distributions that can be very closely fit by an appropriate normal curve.

For example;

- Heights
- Weights
- Some other physical characteristics of humans and animals
- Measurement errors in scientific experiments …..

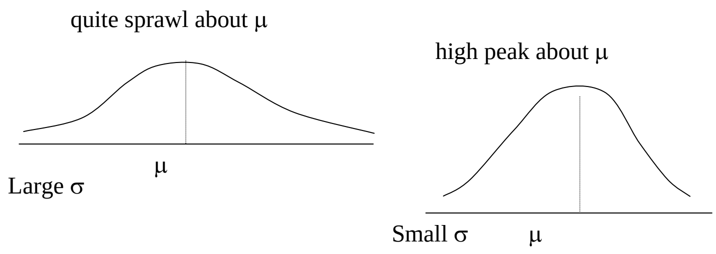
Normal Distribution

- Bell-shaped
- Symmetric about mean
- Continuous
- Never touches the x-axis
- Total area under curve is 1.00
- Approximately
  - 68% lies within 1 standard deviation of the mean,
  - 95% within 2 standard deviations, and
  - 99.7% within 3 standard deviations of the mean.

  this is called as Empirical Rule.

- Data values represented by x which has mean μ and standard deviation σ.

Normal Distribution



Standard Normal Distribution

- Same as a normal distribution, but also ...
  - Mean is zero
  - Variance is one
  - Standard Deviation is one
  - Data values represented by z;

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

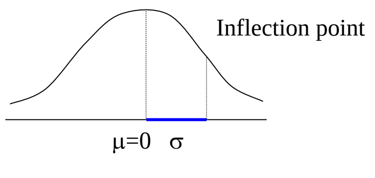
Standard Normal Distribution

Probability density function for standard normal variable Z, with mean 0 and variance 1, is:

$$f(z) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} & -\infty < Z < \infty \\ 0 & \text{o.w.} \end{cases}$$

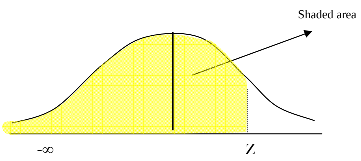
Standard Normal Distribution

- The standard normal density curve, or z curve, is shown below. It is centered at 0 and has inflection points at



Standard Normal Distribution

- Appendix Table 4, is a tabulation of cumulative z curve areas;
  - that is, the table gives areas under the z curve from “0” (origine) to z.
  - To find the area that left of various values, yellow shaded area,



- Entries in this table were obtained by using numerical integration techniques, since the standard normal density function cannot be integrated in a straightforward way.

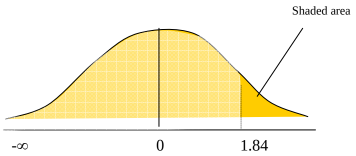
Standard Normal Distribution

Example

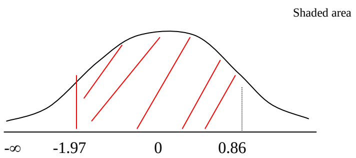
Given a standard normal distribution, find the area under the curve that lies

a) To the right of z = 1.84
b) between z=-1.97 and z=0.86

Standard Normal Distribution



a) The area in figure to the right of z = 1.84 is equal to 1 minus the area in Table to the left of z = 1.84, 1-0.9671 = 0.0329



b) The area in figure between z = -1.97 and z = 0.86 is equal to the the area to the left of z= 0.86 minus the area to the left of z = -1.97. From Table, we find the desired area to be 0.8051-0.0244 = 0.7807

Standard Normal Distribution

Example

- **The probability of values in a standard normal distribution that are less than 1.25 is**

$$\begin{pmatrix}\text{Probability of } z \text{ values satisfying} \\ z < 1.25\end{pmatrix} = \begin{pmatrix}\text{Entry in Table at the intersection of} \\ \text{the } 1.2 \text{ row and } .05 \text{ column}\end{pmatrix}$$
$$= 0.8944$$

- **The probability of values in a standard normal distribution that are less than -0.38 is**

$$\begin{pmatrix}\text{Probability of } z \text{ values satisfying} \\ z < -0.38\end{pmatrix} = 1 - \begin{pmatrix}\text{Entry in Table at the intersection of} \\ \text{the } 0.3 \text{ row and } .08 \text{ column}\end{pmatrix}$$
$$= 1 - 0.648$$
$$= 0.352$$

Standard Normal Distribution

Given a standard normal distribution, find the value of k such that

(a) $P(Z>k) = 0.3015$
(b) $P(k<Z<-0.18)=0.4197$

Normal Distribution

- Any normal curve area can be obtained by first calculating a **"standardized"** limit or limits, and then determining the corresponding area under the **z** curve.
- Let *X* have a normal distribution with parameters μ and λ. Then the standardized variable

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

has a standard normal distribution. This implies that if we form the standardized limits

$$a^* = \frac{a-\mu}{\sigma} \qquad \mathrm{b}^* = \frac{b-\mu}{\sigma}$$

Normal Distribution

- Then

$$\left\{\begin{array}{l}\text{Probability of x values satisfying}\\ a < x < b\end{array}\right\} = \left\{\begin{array}{c}\text{probabilityof z values satisfying}\\ a^* < z < b^*\end{array}\right\}$$

$$\left\{\begin{array}{l}\text{Probability of x values satisfying}\\ x < a\end{array}\right\} = \left\{\begin{array}{c}\text{probability of z values satisfying}\\ z < a^*\end{array}\right\}$$

$$\left\{\begin{array}{l}\text{Probability of x values satisfying}\\ x > b\end{array}\right\} = \left\{\begin{array}{c}\text{probability of z values satisfying}\\ z > b^*\end{array}\right\}$$

Normal Distribution

Example

Given a random variable X having a normal distribution with μ=50, and σ=10, find the probability that X assumes a value between 45 and 62.

$$z_1 = \frac{45-50}{10} = -0.5 \quad \text{and} \quad z_2 = \frac{62-50}{10} = 1.2$$

$$\begin{aligned} P(45 < X < 62) = P(-0.5 < Z < 1.2) \\ &= \mathrm{P(Z < 1.2) - P(Z < -0.5)} \\ &= \mathrm{P(Z < 1.2) - (1 - P(Z < 0.5))} \\ &= 0.8894 - (1 - 0.6915) \\ &= 0.8894 - 0.3085 \ = 0.5764 \end{aligned}$$

Normal Distribution

Example

Given that X has a normal distribution with μ=300, and σ=50, find the probability that X assumes a value greater than 362.

$$P(Z > 362) = P(Z > 1.24) = 1 - P(Z < 1.24)$$
$$= 1 - 0.8925$$
$$= 0.1075$$

Normal Distribution

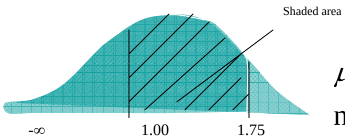
- The reaction time for an in-traffic response to a brake signal from standard brake lights can be modeled with a normal distribution having parameters

$$\mu = 1.25 \text{sec. and } \sigma = .46 \text{ sec.}$$

- In the long run, what is the probability of reaction times that will be between 1.00 sec. and 1.75 sec? Let X denote reaction time. The standardized limits are

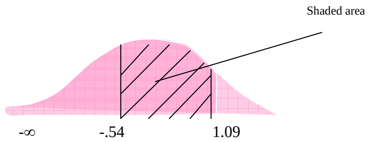
$$\frac{1.00 - 1.25}{.46} = -.54, \quad \frac{1.75 - 1.25}{.46} = 1.09$$

Normal Distribution



$\mu = 1.25, \sigma = .46$

normal dist.



$\mu = 0, \sigma = 1$

z curve

Normal Distribution

If 2 sec is viewed as a critically long reaction time, what is the probability of reaction times that exceed this value?

- P(X>2)=?

$$P(X > 2) = P(\frac{x-\mu}{\sigma} > \frac{2-1.25}{0.46})$$
$$= \mathrm{P}(\mathrm{Z} > 1.63) = 1 - \mathrm{P}(\mathrm{Z} < 1.63)$$
$$= 1 - 0.9484 = 0.0516$$

Normal Distribution

Example:

The amount of distilled water dispensed by a certain machine has normal distribution with

$$\mu = 64 \text{ oz and } \sigma = .78 \text{ oz.}$$

- What container size **c** will ensure that overflow occurs only .5% of the time?

- Let *X* denote the amount of water dispensed. $-\infty$ and c is $.995$.
- The cumulative area under curve above between
- That is, c is the 99.5 th percentile of this normal distribution.

Normal Distribution

Example

Given a normal distribution with $\mu = 40$, and $\sigma = 6$, find the value of X that has

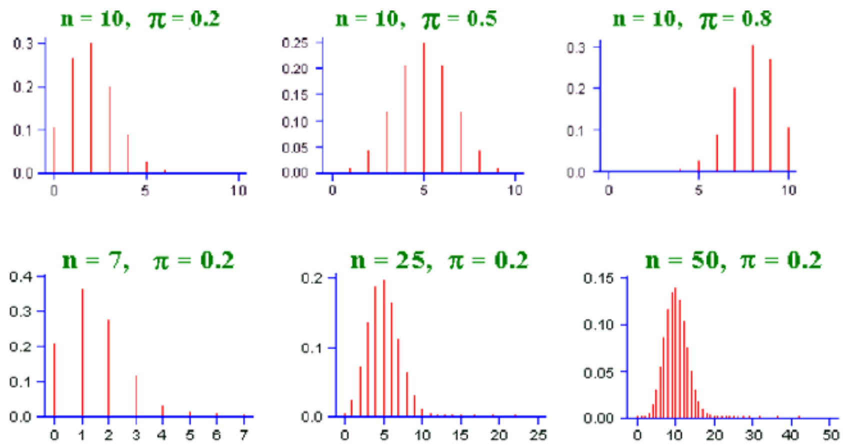
(a) 45% of the area to the left,

(b) 14% of the area to the right

The Normal Approximation to the Binomial

- Computing binomial probabilities using the binomial mass function can be difficult for large *n*.
- If tables are used to compute binomial probabilities, calculations typically are only given for selected values of *n* <= 50 and for selected values of π.
- If n is quite large or if the binomial applet is not available, the normal distribution can be used to approximate the binomial distribution.

The Comparison of Binomial and the Normal Distributions



The Normal Approximation to the Binomial

- For large *n* (say n > 20) and π not too near 0 or 1 ($0.05 < \pi < 0.95$) the distribution approximately follows the Normal distribution.
- This can be used to find binomial probabilities.
- If X ~ binomial (n, π) where n > 20 and $0.05 < \pi < 0.95$ then approximately X has the Normal distribution with mean
- E(X)= $\mu = n\pi$

$$\sigma = \sqrt{n\pi(1-\pi)}$$

$$\text{so } z = \frac{x - n\pi}{\sqrt{n\pi(1-\pi)}} \text{ is approximately N(0,1).}$$

Continuity Correction and Accuracy

- For accurate values for binomial proportions,
  - either use computer software to do exact calculations or
  - if *n* is not very large, the proportion calculation can be improved by using the **continuity correction**.
- This method considers that each whole number occupies the interval from 0.5 below to 0.5 above it.
- When an outcome X needs to be included in the probability calculation,
  - the normal approximation uses the interval from (X-0.5) to (X+0.5). This is illustrated in the following example.

The Normal Approximation to the Binomial

- Example:

In a particular faculty 60% of students are men and 40% are women. In a random sample of 50 students what is the probability that more than half are women?

- Let X= number of women in the sample.

- Assume X has the binomial distribution with
  - n = 50 and π = 0.4. Then E(X) = n π = 50 x 0.4 = 20
  - nπ(1- π) = 50 x 0.4 x 0.6 = 12 , $\sigma = \sqrt{npq} = \sqrt{12} = 3.44$
  - so approximately X ~ N(20, 3.44).
  - We need to find P(X > 25). Note - **not** P(X >= 25).

$$if\ \ \text{x} > 25 \text{ then z} = \frac{25-20}{\sqrt{12}} = 1.44$$

  - P(X > 25) = P(Z > 1.44) = 1 - P(Z < 1.44) = 1 - 0.9251 = 0.075

The Normal Approximation to the Binomial

- The exact answer calculated from binomial probabilities is P(X>25) = P(X=26) + P(X=27) + ... + P(X=50) = 0.0573

- The approximate probability, using the **continuity correction**, is

$$p(x > 25) = p(z > \frac{25.5-20}{\sqrt{12}}) = p(z > 1.5877)$$

Using entry in row 1.5 and column .08 in Table 4

$p(z > 1.5877) = 0.5 - .4429 = 0.0571$

- 0.0571 which is a much better approximation to the exact value of 0.0573

- (The value 25.5 was chosen as the outcome 25 was **not** to be included but the outcomes 26, 27, 50 **were** to be included in the calculation.)

The Normal Approximation to the Binomial

•Similarly, if the example requires the probability that less than 18 students were women,

• the continuity correction would require the following calculation:

$$p(x < 18) = p(z < \frac{17.5 - 20}{\sqrt{12}}) = p(z < -2.5/3.44) =$$
$$p(z < -.726) = .2358 = \sim 23.6\%$$

The Normal Approximation to the Binomial

Example X has binomial distribution with p = 0.4 and n = 15.

a) P(X=4) = ?

Using binomial distribution P(X=4) = 0.1268

When normal approximation to binomial distribution is used;

$$\mu = np = (15)(0.4) = 6 \quad \text{and} \quad \sigma^2 = npq = (15)(0.4)(0.6) = 3.6 \quad \text{and} \quad \sigma = \sqrt{3.6} = 1.897$$

$$z_1 = \frac{(4 - 0.5) - 6}{1.897} = -1.32 \qquad \text{and } z_2 = \frac{(4 + 0.5) - 6}{1.897} = -0.79$$

$$\begin{aligned} P(X = 4) &= P(-1.32 < Z < -0.79) \\ &= \text{P(Z<-0.79)-P(Z<-1.32)} \\ &= 0.2148\text{-}0.0934 \\ &= 0.1214 \end{aligned}$$

The Normal Approximation to the Binomial

b) P(7 ≤ X ≤ 9) =?

When binomial distribution is used, P(7 ≤ X ≤ 9) = 0.3546

Using normal approximation to binomial distribution

$$z_1 = \frac{(7-0.5)-6}{1.897} = 0.26 \quad \text{and} \quad z_2 = \frac{(9+0.5)-6}{1.897} = 1.85$$

$$\begin{aligned} P(7 \le X \le 9) &= P(0.26 < Z < 1.85) \\ &= \mathrm{P}(\mathrm{Z} < 1.85) - \mathrm{P}(\mathrm{Z} < 0.26) \\ &= 0.9678 - 0.6026 \\ &= 0.3652 \end{aligned}$$

References

- Walpole, Myers, Myers, Ye, (2002),
  - *Probability & Statistics for Engineers & Scientists*
- Dengiz, B., (2004),
  - *Lecture Notes on Probability,* http://w3.gazi.edu.tr/web/bdengiz
- Hines, Montgomery, (1990),
  - *Probability & Statistics in Engineering* & Management *Science*