

VI. Continuous Probability Distributions

A. An Important Definition (reminder)

Continuous Random Variable - a numerical description of the outcome of an experiment whose outcome can assume any numerical value in an interval or collection of intervals.

B. Some Important Continuous Probability Distributions

1. (Continuous) Uniform Probability - expresses the likelihoods of outcomes for a continuous random variable x for which all possible outcomes are equally likely. This distribution function is given by

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

where a = minimum possible value of random variable x
 b = maximum possible value of random variable x

Characteristics of a (Continuous) Uniform probability distribution are:

- the random variable can assume any value within a range
- all possible values within this range are equally likely
- no value outside of this range can occur

The mean (expected value) is

$$E(x) = \frac{a+b}{2}$$

and the variance is

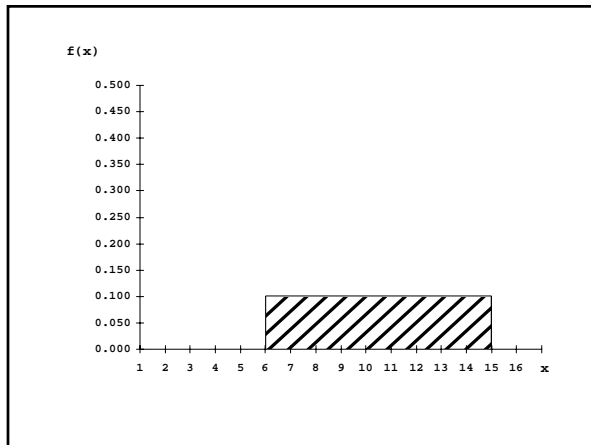
$$\sigma^2 = \frac{(b-a)^2}{12}$$

but the probability is given by

$$P(c \leq x \leq d) = \begin{cases} \frac{d - c}{b - a} & \text{for } a \leq c, d \leq b \\ 0 & \text{otherwise} \end{cases}$$

Why isn't $f(x) = P(x)$? Think about the graph of the probability distribution function for a (continuous) uniform random variable with a range of 6 - 15, i.e.,

$$f(x) = \begin{cases} \frac{1}{9} & \text{for } 6 \leq x \leq 15 \\ 0 & \text{otherwise} \end{cases}$$



What is the area under the probability distribution function at $f(x)$?

Consider $x = 8$. We have that $f(8) = 0.111$, but $P(8)$ is the product of the height of $f(x)$ (the probability distribution) and the width at $x = 8$. Thus

$$P(x) = 0.111(0.00) = 0.00$$

This leads to two conclusions about working with probability distribution functions for continuous random variables:

- $P(x) = 0$ for any single value of x
- we can only talk meaningfully about the probability over a range of values

So how can we calculate probabilities for continuous random variables over ranges?

To find $P(a \leq x \leq b)$ we integrate the probability distribution function $f(x)$ from a to b , i.e.,

$$\int_a^b f(x) dx$$

The mean (expected value) is

$$E(x) = \mu = \int_{\text{all } x} xf(x) dx$$

and the variance is

$$\sigma^2 = \int_{\text{all } x} (x - \mu)^2 f(x) dx$$

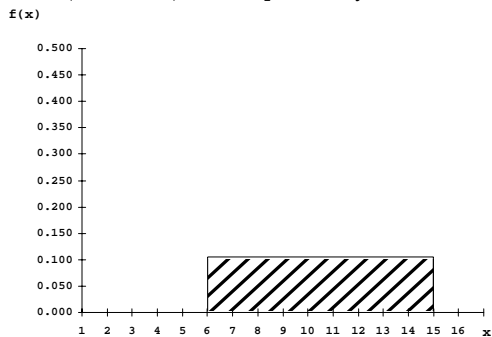
Example - Suppose that some random variable x can take on any value between 6 and 15 with equal probability, and can take on no value outside of this range. What is the probability that x is between 7 and 10?

$$P(7 \leq x \leq 10) = \frac{10 - 7}{15 - 6} = \frac{3}{9} = 0.333$$

What is the probability that x is between 4 and 8?

$$\begin{aligned} P(4 \leq x \leq 8) &= P(4 \leq x < 6) + P(6 \leq x \leq 8) \\ &= \frac{0}{15 - 6} + \frac{8 - 6}{15 - 6} = \frac{0}{9} + \frac{2}{9} = 0.222 \end{aligned}$$

Why didn't we have to integrate to find probabilities for the (continuous) uniform probability distribution?



2. Exponential Probability Distribution - expresses the likelihoods of outcomes for a continuous random variable x that represents the amount of time or space that passes between consecutive occurrences of a Poisson random variable. This distribution function is given by

$$f(x) = \frac{1}{\mu} e^{-x/\mu} \text{ for } x \geq 0, \mu > 0$$

where μ = mean time or space between consecutive occurrences

$e = 2.71828\dots$

Characteristics of an Exponential probability distribution are:

- the random variable can assume any positive value
- the random variable represents the amount of the interval (time, space, etc.) that passes between consecutive occurrences of a Poisson event

The mean (expected value) is

$$E(x) = \mu$$

and the variance is

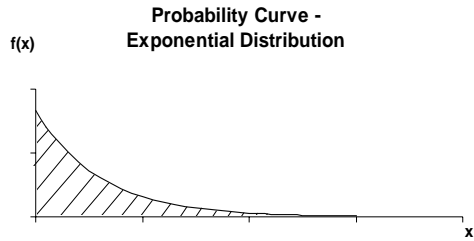
$$\sigma^2 = \mu$$

but the probability is given by

$$P(x \leq b) = P(0 \leq x \leq b) = \int_0^b f(x) dx$$
$$= \int_0^b \frac{1}{\mu} e^{-x/\mu} dx = 1 - e^{-b/\mu}$$

Fortunately, the integration for this probability distribution function is relatively easy and yields a closed-form result! This means we will not have to do the integration to find probabilities for an exponential random variable?

Generically, the exponential probability distribution curve looks like this:



Example: suppose that the time that passes between customers arriving at an ATM is exponentially distributed with a mean of 6 minutes. What is the probability that no more 0.05 hours will pass between arrivals of the next two customers?

Call random variable x the amount of time (in minutes) that passes between consecutive arrivals. We have that

$\mu = 6.0$ minutes between consecutive customers

and

$x = 0.05$ hours = $0.05(60)$ minutes = 3.0 minutes

so

$$P(x \leq 3.0) = 1 - e^{-3.0/6.0} = 0.3935$$

OR IN HOURS

$\mu = 6.0$ minutes = 0.10 hours between consecutive customers

and

$x = 0.05$ hours

so

$$P(x \leq 0.05) = 1 - e^{-0.05/0.10} = 0.3935$$

To illustrate the relationship between the Exponential distribution and the Poisson distribution, consider the following example:

Suppose that an ATM averages 10 customer arrivals per hour, and that these arrivals are Poisson distributed. What is the probability that no more than three minutes will pass between arrivals of the next two customers?

Call random variable x the amount of time (in minutes) that passes between consecutive arrivals. We have that

$$\mu = \frac{60}{10} = 6.0 \text{ minutes between customers}$$

so

$$P(x \leq 3.0) = 1 - e^{-3.0/6.0} = 0.3935$$

Again, this same problem could be done in terms of hours (instead of minutes). Call random variable y the amount of time (in hours) that passes between consecutive arrivals.

We have that

$$\mu = \frac{1}{10} = 0.10 \text{ hours between customers}$$

so

$$P(x \leq 0.05) = 1 - e^{-0.05/0.10} = 0.3935$$

O.K. - we can find probabilities of the form $P(x \leq b)$.
What about problems of the form $P(x \geq b)$ or $P(a \leq x \leq b)$?

We can now use our original result for $P(x \leq b)$.
Remember that $P(x \geq b) = 1 - P(x \leq b)$, so

$$P(x \geq b) = 1 - P(x \leq b) = 1 - \left(1 - e^{-b/\mu}\right) = e^{-b/\mu}$$

Similarly, $P(a \leq x \leq b) = P(x \leq b) - P(x \leq a)$, so

$$P(a \leq x \leq b) = P(x \leq b) - P(x \leq a) \\ = \left(1 - e^{-b/\mu}\right) - \left(1 - e^{-a/\mu}\right) = e^{-a/\mu} - e^{-b/\mu}$$

Note that Exponential Probabilities Tables do exist for certain values of μ , but these are not provided in your textbook. Also note that

$$E(x) = \mu$$

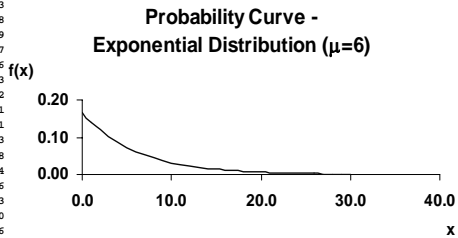
and

$$\text{Var}(x) = \sigma^2 = \mu$$

for the exponential distribution.

x	$f(x)$
0.0	0.166667
1.0	0.141080
2.0	0.119422
3.0	0.101088
4.0	0.085570
5.0	0.072433
6.0	0.061313
7.0	0.051901
8.0	0.043933
9.0	0.037168
10.0	0.031479
11.0	0.026647
12.0	0.022556
13.0	0.019093
14.0	0.016162
15.0	0.013681
16.0	0.011581
17.0	0.009803
18.0	0.008298
19.0	0.007024
20.0	0.005946
21.0	0.005033
22.0	0.004260
23.0	0.003606
24.0	0.003053
25.0	0.002584

The probability curve and distribution for the previous problem (a Poisson distributed random variable with a mean of $\mu = 7$ customers per hour) look like this:



3. Normal Probability Distribution - expresses the likelihoods of outcomes for a continuous random variable x with a particular symmetric and unimodal distribution. This distribution function is given by

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

where μ = mean

σ = standard deviation

π = 3.14159

e = 2.71828

but the probability is given by

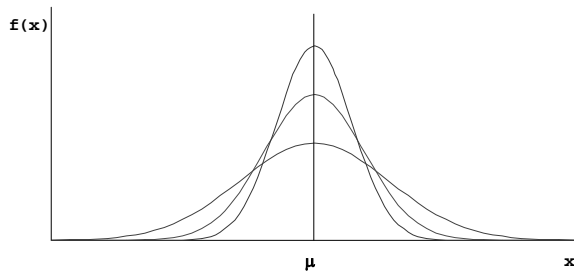
$$P(a \leq x \leq b) = \int_a^b f(x) dx$$
$$= \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

This looks like a difficult integration problem! Will I have to integrate this function every time I want to calculate probabilities for some normal random variable?

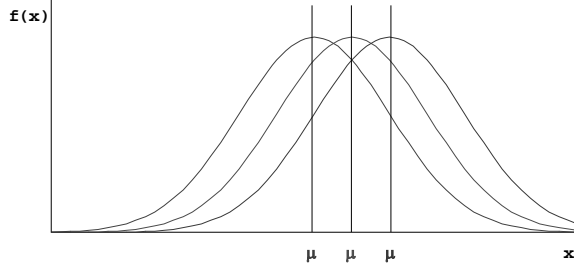
Characteristics of the normal probability distribution are:

- there are an infinite number of normal distributions, each defined by their unique combination of the mean μ and standard deviation σ
- μ determines the central location and σ determines the spread or width
- the distribution is symmetric about μ
- it is unimodal
- $\mu = M_d = M_o$
- it is asymptotic with respect to the horizontal axis
- the area under the curve is 1.0
- it is neither platykurtic nor leptokurtic
- it follows the empirical rule

Normal distributions with the same mean but different standard deviations:



Normal distributions with the same standard deviation but different means:



The Standard Normal Probability Distribution - the probability distribution associated with any normal random variable that has $\mu = 0$ and $\sigma = 1$.

There are tables that give the results of the integration

$$P(a \leq x \leq b) = \int_a^b f(x) dx$$

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

for the standard normal random variable (Table 1, Appendix A).

The Cumulative Standard Normal Distribution (Appendix B, Table 1)

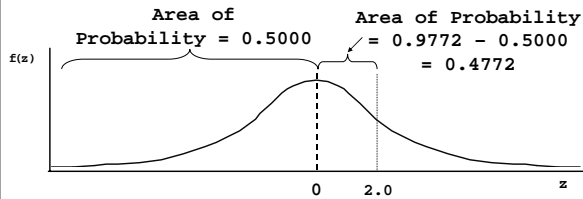
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.0	0.0013	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018
-2.9	0.0019	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025
-2.8	0.0026	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034
-2.7	0.0035	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045
-2.6	0.0047	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060
-2.5	0.0062	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080
-2.4	0.0082	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104
-2.3	0.0107	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136
-2.2	0.0139	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174
-2.1	0.0179	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222
-2.0	0.0228	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281
-1.9	0.0287	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351
-1.8	0.0359	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436
-1.7	0.0446	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537
-1.6	0.0548	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655
-1.5	0.0668	0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793
-1.4	0.0808	0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951
-1.3	0.0968	0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131
-1.2	0.1151	0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335
-1.1	0.1357	0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562
-1.0	0.1587	0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814
-0.9	0.1841	0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090
-0.8	0.2119	0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389
-0.7	0.2420	0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709
-0.6	0.2743	0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050
-0.5	0.3085	0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409
-0.4	0.3446	0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783
-0.3	0.3821	0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168
-0.2	0.4207	0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562
-0.1	0.4602	0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359

Again, looking at a small part of the Cumulative Standard Normal Probability Distribution Table, we find the probability that a standard normal random variable z is between $-\infty$ and 2.00?

z	0.00	0.01	0.02	0.03	0.04
:	:	:	:	:	:
:	:	:	:	:	:
1.3	0.9032	0.9049	0.9066	0.9082	0.9099
1.4	0.9192	0.9207	0.9222	0.9236	0.9251
1.5	0.9332	0.9345	0.9357	0.9370	0.9382
1.6	0.9452	0.9463	0.9474	0.9484	0.9495
1.7	0.9554	0.9564	0.9573	0.9582	0.9591
1.8	0.9641	0.9649	0.9656	0.9664	0.9671
1.9	0.9713	0.9719	0.9726	0.9732	0.9738
2.0	0.9772	0.9778	0.9783	0.9788	0.9793
2.1	0.9821	0.9826	0.9830	0.9834	0.9838

Example: for a standard normal random variable z , what is the probability that z is between 0 and 2.0?

Area of Probability = 0.9772

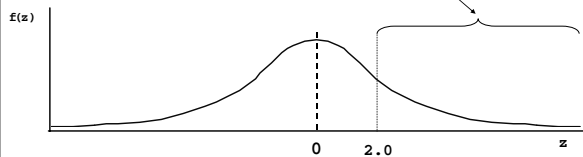


$$P(0 \leq z \leq 2) = P(-\infty \leq z \leq 2) - P(-\infty \leq z \leq 0)$$

$$= 0.9772 - 0.5000 = 0.4772$$

What is the probability that z is at least 2.0?

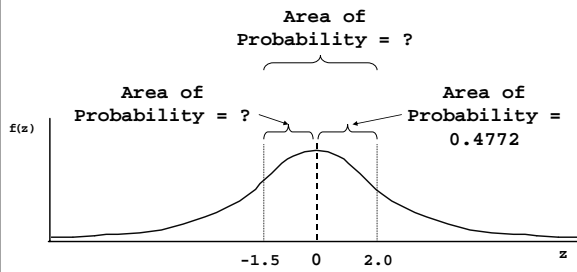
Area of Probability =
1.0000 - 0.9772 = 0.0228



$$P(z \geq 2) = P(-\infty \leq z \leq \infty) - P(-\infty \leq z \leq 2)$$

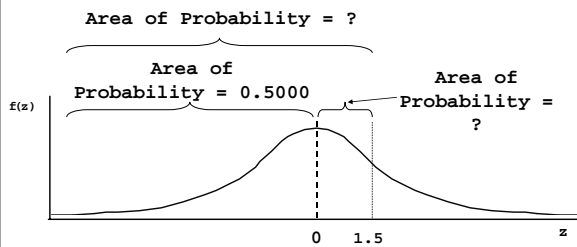
$$= 1.0000 - 0.9772 = 0.0228$$

What is the probability that z is between -1.5 and 2.0 ?



We need to find the probability that z is between -1.5 and 0.0 !

By symmetry of the standard normal distribution, we have equivalently

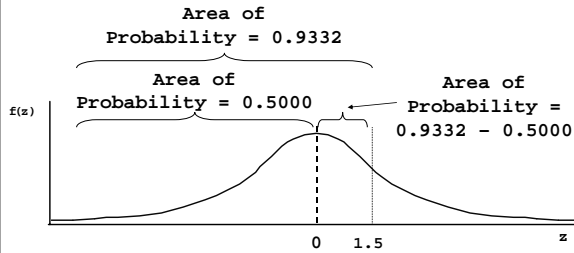


Note that we have simply inverted the previous problem so that we can use the cumulative standard normal distribution table!

Again, looking at a small part of the Cumulative Standard Normal Probability Distribution Table, we find the probability that a standard normal random variable z is between $-\infty$ and 1.50 ?

z	0.00	0.01	0.02	0.03	0.04
:	:	:	:	:	:
:	:	:	:	:	:
1.3	0.9032	0.9049	0.9066	0.9082	0.9099
1.4	0.9192	0.9207	0.9222	0.9236	0.9251
1.5	0.9332	0.9345	0.9357	0.9370	0.9382
1.6	0.9452	0.9463	0.9474	0.9484	0.9495
1.7	0.9554	0.9564	0.9573	0.9582	0.9591
1.8	0.9641	0.9649	0.9656	0.9664	0.9671
1.9	0.9713	0.9719	0.9726	0.9732	0.9738
2.0	0.9772	0.9778	0.9783	0.9788	0.9793
2.1	0.9821	0.9826	0.9830	0.9834	0.9838

...so by symmetry of the standard normal distribution, we have equivalently

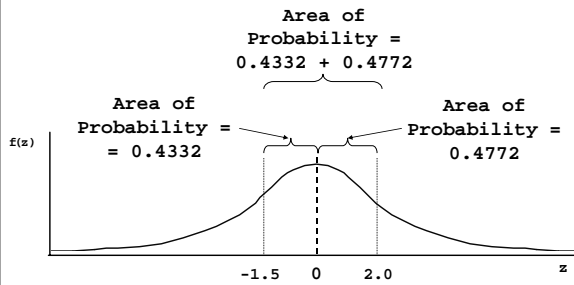


$$P(-1.5 \leq z \leq 0) = P(0 \leq z \leq 1.5)$$

$$= [P(-\infty \leq z \leq 1.5) - P(-\infty \leq z \leq 0)]$$

$$= [0.9332 - 0.5000] = 0.4332$$

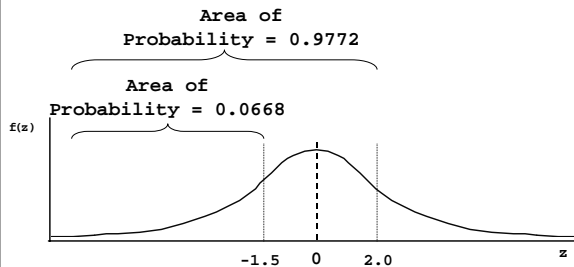
What is the probability that z is between -1.5 and 2.0?



$$P(-1.5 \leq z \leq 2.0) = P(-1.5 \leq z \leq 0) + P(0 \leq z \leq 2.0)$$

$$= 0.4332 + 0.4772 = 0.9104$$

Notice we could find the probability that z is between -1.5 and 2.0 another way!



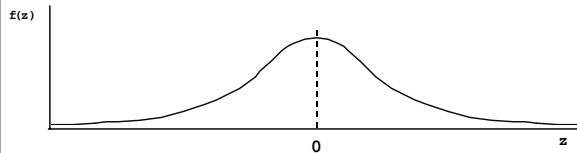
$$P(-1.5 \leq z \leq 2.0) = P(-\infty \leq z \leq 2.0) - P(-\infty \leq z \leq -1.5)$$

$$= 0.9772 - 0.0668 = 0.9104$$

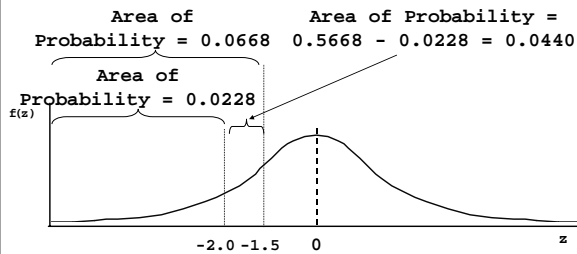
There are often multiple ways to use the Cumulative Standard Normal Probability Distribution Table to find the probability that a standard normal random variable z is between two given values!

How do you decide which to use?

- Do what you understand (make yourself comfortable)
- and
- **DRAW THE PICTURE!!!**



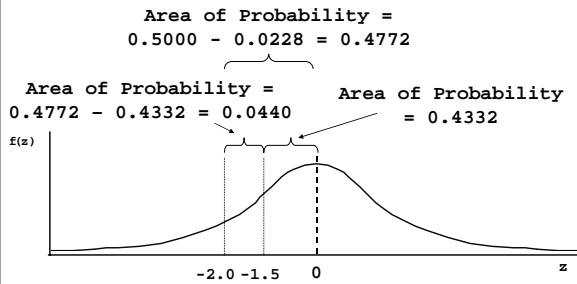
What is the probability that z is between -1.5 and -2.0?



$$P(-2.0 \leq z \leq -1.5) = P(-\infty \leq z \leq -1.5) - P(-\infty \leq z \leq -2.0)$$

$$= 0.0668 - 0.0228 = 0.0440$$

What is the probability that z is between -1.5 and -2.0?

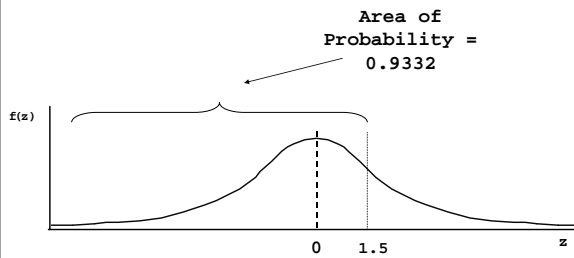


$$P(-2.0 \leq z \leq -1.5) = P(-2.0 \leq z \leq 0.0) - P(-1.5 \leq z \leq 0.0)$$

$$= [P(-\infty \leq z \leq 0.0) - P(-\infty \leq z \leq -2.0)] - 0.4332$$

$$= [0.5000 - 0.0228] - 0.4332 = 0.0440$$

What is the probability that z is *exactly* 1.5?



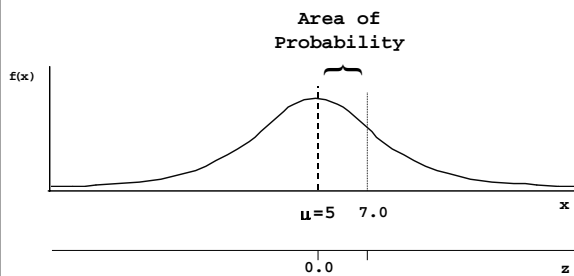
$$\begin{aligned}
 P(z = 1.5) &= P(1.5 \leq z \leq 1.5) \\
 &= P(-\infty \leq z \leq 1.5) - P(-\infty \leq z \leq -1.5) \\
 &= 0.9332 - 0.9332 = 0.0000 \quad (\text{why?})
 \end{aligned}$$

z-Transformation - mathematical means by which any normal random variable with a mean μ and standard deviation σ can be converted into a standard normal random variable.

- to make the mean equal to 0, we simply subtract μ from each observation in the population
 - to then make the standard deviation equal to 1, we divide the results in the first step by σ
- The resulting transformation is given by

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

Example: for a normal random variable x with a mean of 5 and a standard deviation of 3, what is the probability that x is between 5.0 and 7.0?



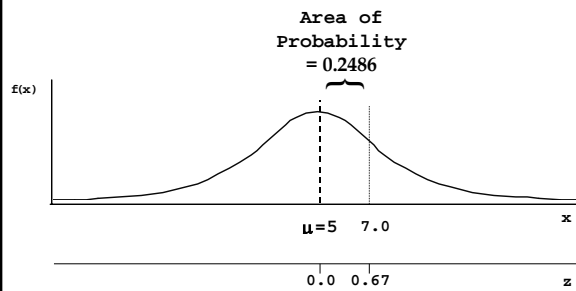
Using the z-transformation, we can restate the problem in the following manner:

$$P(5.0 \leq x \leq 7.0) = P\left(\frac{5.0 - 5.0}{3.0} \leq \frac{x - \mu}{\sigma} \leq \frac{7.0 - 5.0}{3.0}\right) \\ = P(0.0 \leq z \leq 0.67)$$

then use the standard normal probability table to find the ultimate answer:

$$P(0.0 \leq z \leq 0.67) = 0.2486$$

which graphically looks like this:



Why is the normal probability distribution considered so important?

- many random variables are naturally normally distributed
- many distributions, such as the Poisson and the binomial, can be approximated by the normal distribution
- the distribution of many statistics, such as the sample mean and the sample proportion, are approximately normally distributed if the sample is sufficiently large (Central Limit Theorem)

