**Lab 5**

**Normal Distribution**

* Used with continunous random variables. 
* Characterized by 2 parameters: *μ* = mean

*σ* = standard deviation

* Possible values: any possible value
* Shape of distribution: symmetric, bell shaped
* Assumptions: that the data is symmetric, bell shaped
* Mean : *μ*
* Variance: *σ 2*
* Standard Deviation: *σ*
* Standard normal distribution: *μ* = \_0\_ and *σ* = \_1\_ . Typically, the letter *Z* denotes variables that follow a standard normal distribution. Probabilities associated with the standard normal distribution are tabulated in many books.
* Transforming from an arbitrary normal random variable to a standard normal random variable:

P(X≤x)=P(Z≤ (x-)/)

X=x

Z=(x-)/

**Calculator in Stata**

To do any calculations in Stata you have to use the command “di”. Let’s say you have a normal population with mean 115 and standard deviation 3. You have a data point from this population which is 120. You want to get the z-score from Stata, the following command gives you what you want:

di (120-115)/3

1.666667

**Calculating Normal Probabilities with Stata**

Like binomial probabilities, Stata can be used to calculate probabilities associated with the standard normal distribution using the normal command. To calculate the probability that a standard normal random variable *Z* is less than some value k (i.e. P(Z<k)), type

di normal(k)

If you want the probability that *Z* is greater than some value k (i.e. P(Z>k)), then make use of the fact that

P(Z<k)+ P(Z>k)=1

and use the command

di 1-normal(k)

If you have a random variable, *X*, that follows an arbitrary normal distribution with mean *μ* and standard deviation *σ*, you must first transform X into a standard normal random variable Z before using the normal command.

To find the value Z of the standard normal distribution that cuts off an area of p in the left tail of the standard normal curve use:

di invnormal(p)

**Helpful Hint:** When calculating normal probabilities, ALWAYS draw the distribution. Visualization helps and it will be useful for future concepts.

**Question 1**

Suppose that *X* is a random variable that represents the height of women ages 18-74 in the US. Assume that *X* is normally distributed with a mean of 63.9 inches and a standard deviation of 2.6 inches.

1. What is the probability that a randomly selected woman is less than 60 inches tall?

di normal(z)

P(X<60)

0

-1.5

convert to std normal

(X-*μ*)/*σ* = Z

P(X<60)= P[(X-*μ*)/*σ* < (60 – 63.9)/2.6]

= P(Z< -1.5)= 0.06680

= .0668

63.9

60

di normal(-1.5)= 0.06680

1. What is the probability that a randomly selected woman is greater than 68 inches tall?

di 1-normal(z)

P(X>=68)

0

1.587

convert to std normal

(X-*μ*)/*σ* = Z

P(X>=68)= P[(X-*μ*)/*σ* >=(68– 63.9)/2.6]

= P(Z>= 1.577)= 0.0570

= .0574

63.9

68

1. What is the probability that a randomly selected woman is between 60 and 68 inches tall?

P(60<=X<=68)

0

1.587

convert to std normal

(X-*μ*)/*σ* = Z

P(60<=X<=68)

=P[(60 – 63.9)/2.6 <=(X-*μ*)/*σ* <=68– 63.9)/2.6]

= P(-1.5<=Z<= 1.58)= 0.87613

kmm

= 1-P(Z<= -1.5) – P(X>=1.58)

=1-.0668-.0574=.876

63.9

68

60

-1.5

di normal(1.58)-normal(-1.5)=0.87613

1. What height would be the 95th percentile?

di invnormal(0.95)

0

1.587

convert to the X normal dstbn

X=

P(Z>1.645)= P[Z\**σ +μ* >1.645\*2.6+69.3]

= P(X>= 1.577)

= .68.11

0

1.645

Therefore, 95 % of the population is less than 68.11 inches.

1. What is the probability that among 5 women selected at random from this population, exactly one will be between 60 and 68 inches tall?

Note that that this follows a binomial distribution (2 events: either the woman is or is not in this range, independence, and same probability of event). Therefore Y is distributed binomial with n=5 and p= .876.

P(X=1) = .00104

di binomial(5,1,0.876)-binomial(5,0,0.876)=0.00104

**Question 2**

Among females in the United States between 18 and 74 years of age, diastolic blood pressure is normally distributed with mean μ=77 mmHg and standard deviation σ = 11.6 mmHg.

1. What is the probability that a randomly selected woman has a diastolic blood pressure less than 60 mmHg?

P(X<60)

0

-1.47

convert to std normal

(X-*μ*)/*σ* = Z

P(X<60)= P[(X-*μ*)/*σ* < (60 – 77)/11.6]

= P(Z< -1.47)= 0.07078

= .0713

77

60

1. What is the probability that she has a diastolic blood pressure greater than 90 mmHg?

P(X>90)

0

1.127

convert to std normal

(X-*μ*)/*σ* = Z

P(X>90)= P[(X-*μ*)/*σ* >(90– 77)/11.6]

= P(Z> 1.121)=0.1311

= .1311

63.9

90

1. What is the probability that she has a diastolic blood pressure between 60 and 90 mmHg?

P(60<=X<=68)

0

1.587

convert to std normal

(X-*μ*)/*σ* = Z

P(60<=X<=90)

=P[(60 – 77)/11.6 <(X-*μ*)/*σ* <(90– 77)/11.6]

= P(-1.47<=Z<= 1.12)=0 .7978

= 1-P(Z<= -1.47) – P(X>=1.12)

=1-.1311-.0713=.7976

77

90

60

-1.5

**Theoretical Distributions: Review Table**

|  |  |  |
| --- | --- | --- |
|  | **Binomial** | **Normal** |
| Parameters | **n, p** | *μ*, *σ* |
| Possible Values | **0,1,…n** | **all real values** |
| Mean | **n\*p** | *μ* |
| Standard Deviation | [n\*p\*(1-p)]1/2 | *σ* |