**Business Statistics – Probability & Probability Distributions**

**Continuous Probability Distributions**

|  |  |
| --- | --- |
|  | Introduction to  Probability Definitions: Frequency, Classical, Subjective  Distributions: Discrete, Continuous, Sampling  Central Limit Theorem, CLT |
|  | Discrete Distributions, Empirical, Binomial, Poisson |
| ► | Continuous Distributions, Empirical, Normal, Exponential |
|  | Bayesian Probability Analysis |

|  |
| --- |
| **For Discrete Distributions.**  For X~P(X), properties are (1) 0<=P(X)<=1; and (2) P(X)=1.  Expectation of X, E[X] = X\*P(X) = .  **For Continuous Distributions.**  If X~f(X), then ∫ f(X) dX = 1,  E[X]= =   P[a<X<b] = |

|  |  |
| --- | --- |
| **Empirical Distribution**  Uniform, X~U(a,b)  Density function, f(X)= 1/(b-a), 0<a<X<b  E[X]=(a+b)/2  Var(X)=(b-a)2/12  P[X<c]=(c-a)/(b-a), c is a constant, a≤c≤b | **a=1**  **b=6** |
| **Exponential Distribution**  X~Exp( l ), l > 0  Density function, f(X)= l exp(-lX), X>0  E[X]=1/l  Var[X]= 1/l2  P[X<c]=1 – exp(-lX), c is a constant, c>0 | **l = 2** |
| **Normal Distribution**  X~N( m, s2 ),  Density function, –∞<X<+∞  f(X)= 1/( ssqrt(2p))exp(-1/2((X- m)/ s)2)  E[X]= m  Var[X]= s2  P[X<c]=, c is a constant | **m = 5**  **s2 = 25** |

|  |  |  |
| --- | --- | --- |
|  | **Normal Probability Distribution** |  |

|  |
| --- |
| **Normal Probability Distribution. X ~ N(  , 2 )** |
| Let the random variable, X, follow the Normal probability distribution  with a mean of  and a variance of 2 .  E[X] =  . Var[X] = 2  This is represented by X ~ N(  , 2 ) .  Probability measure of X is the area under the Normal distribution. |

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  |   |
|  |  |  |  |
| X  X X    X~N(  ,     |  |  | | --- | --- | |  = Probability  X = Normal Random Variable  X = X1- = Normal Variate   = Population Mean  2 = Population Variance   = Population Standard Deviation | | | P[X < X ] =   P[X < X ] =   P[X <  ] = 0.5 , thus, =X0.5 | P[X > X ] =   P[X > X ] =   P[X >  ] = 0.5 | | P[ X < X < X ] = 2  P[X =  ] = 0  P[X = X ] = 0  P[X = X ] = 0 | | | | | |

|  |  |  |
| --- | --- | --- |
|  | **Standard Normal Probability Distribution** |  |

|  |
| --- |
| **Standard Normal Probability Distribution. Z ~ N( 0 , 1 )** |
| **Consider the transformation, Z = ( X –  ) / sqrt( 2 )** |
| **Consider the transformation, X =  + Z \* sqrt( 2 )** |
| The variable, Z, is called the Standard Normal random variable.  Z follows the Standard Normal probability distribution  with a mean of 0 and a variance of 1 .  This is represented by Z ~ N( 0 , 1 ) . |

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  |   |
|  |  |  |  |
| X  X X    X~N(  ,   | | | |
|   |  |  |  |
|  |  |  |  |
| Z  Z Z  Z~N(  , =1or Z~N(  ,  | | | |

|  |  |
| --- | --- |
|  = Probability  Z = Standard Normal Random Variable  Z = Z1- = Standard Normal Variate   = 0 = Population Mean   = 1 = Population Standard Deviation  2 = 1 = Population Variance | Specifically,P[ Z < Z ] =   P[ Z < Z1- ] = 1-  P[ Z > Z1- ] =   P[ Z <  ] =  , thus, 0=Z0.5  P[ Z > Z1- ] =   P[ Z < Z < Z1- ] = 1-2 |

|  |
| --- |
| Since  and 2 are constants, and Z=(X–)/ note:  E[ (X –  ) /  ] = (E[X] –  ) /  ] = (  –  ) /  ] = 0  Var[ (X –  ) /  ] = Var( X /  ) = Var[X]/ 2 = 2 / 2 = 1  Thus, since Z=(X–)/ , E[Z]=0 and Var[Z]=1.  Therefore, Z ~ N( 0 , 1 ) |

|  |  |  |
| --- | --- | --- |
|  | **Normal Probability Calculations** |  |

|  |  |  |  |
| --- | --- | --- | --- |
|   |  |  |  |
|  |  |  |  |
| Z  Z Z  Z~N(  , 1 | | | |
| X  X X  X~N(  ,   | | | |

Calculations using Excel.

|  |
| --- |
| For Normal Distribution, X~N(  ,    For Normal Probability, “=normdist(x,mean,standard\_dev,cumulative)”  For Normal Variate, “=norminv(probability,mean,standard\_dev)” (will use cumulative probability) |
| For Standard Normal Distribution, Z~N(  , 1  For Standard Normal Probability, “=normsdist(z)” (will return cumulative probability)  For Standard Normal Variate, “=normsinv(probability)” (will use cumulative probability) |

|  |  |
| --- | --- |
| Probability Calculations | normsdist(z)=normdist(x,0,1,1) |
| P[ Z < 0 ] = P[ Z > 0 ] = 0.5 | =normsdist(0) |
| P[ Z < 1 ] = P[ Z > –1 ] ≈ 0.8413 | =normsdist(1), =1–normsdist(–1) |
| P[ Z>1 ] = P[ Z< –1 ] ≈ 1 – 0.8413 ≈ 0.1587 | =1–normsdist(1), =normsdist(–1) |
| P[ 0.2 < Z < 0.5 ] = P[ –0.5 < Z < –0.2 ] | =normsdist(0.5) – normsdist(0.2) |
| ≈ 0.6915 – 0.5793 ≈ 0.1122 | =normsdist(–0.2) – normsdist(–0.5) |
| P[ –0.2 < Z < 0.5 ] = P[ –0.5 < Z < 0.2 ] | =normsdist(0.5) – normsdist(–0.2) |
| ≈ 0.6915 + 0.5793 – 1 ≈ 0.2708 | =normsdist(0.2) – normsdist(–0.5) |
| P[ Z < 1.23 ] ≈ 0.8907 | =normsdist(1.23) |
| P[ Z > 1.23 ] ≈ 1 – 0.8907 ≈ 0.1093 | =1–normsdist(1.23) |
| Note that for continuous distributions, P[Z=1.23]=0. Therefore, P[Z<1.23]≡P[Z<=1.23]≈0.8907 | |
|  |  |
| Variate Calculations. (Percentile Calculations.) | normsinv(probability)=norminv(probability,0,1) |
| P[Z<Z]=0.791, Z≈0.81 | =normsinv(0.791) |
| P[Z>Z]=0.209, Z(1–0.209)=Z≈0.81 | =normsinv(1–.209) |
| P[Z>Z]=1–P[Z<Z]≈1–0.791≈0.209 |  |
| thus, Z in the expression is same as Z791≈0.81 |  |

|  |
| --- |
| **Empirical Rule:**  P[ - < X < +] = P[ –1 < Z < +1]≈0.6826  P[ -2 < X < +2] = P[ –2 < Z < +2]≈0.9544  P[ -3 < X < +3] = P[ –3 < Z < +3]≈0.9974 |

|  |  |  |
| --- | --- | --- |
|  | **Normal Probability Examples** |  |

1. A manufacturer produces bearings with diameters required to be 1.2 centimeters (cm). Because of variability in the production process, bearings will have different diameters. The diameters have a normal distribution with a mean of 1.2 cm and standard deviation of 0.02 cm.

|  |
| --- |
| a. If only diameters in the range of 1.16 to 1.24 cm are acceptable, what proportion of all bearings fall in this acceptable range? ANSWER: The acceptable percentage is 95.45%.  Assume X~N(1.2,0.022), where =1.2 and =0.02.  P[1.16 < X < 1.24]=P[(1.16-1.2)/0.02 < (X-)/ < (1.24-1.2)/0.02 ]  =P[ –2 < Z < +2 ]= P[ Z < +2 ] – P[ Z < –2 ]≈ 0.97725 – 0.02275 ≈ 0.9545  =normsdist(+2)–normsdist(–2) = normdist(1.24,1.2,0.02,1)–normdist(1.16,1.2,0.02,1) ≈ 0.9545 |
| b. What symmetric range about the mean includes 90% of all bearings?  Assume X~N(1.2,0.022), where =1.2 and =0.02.  The 5th and 95th percentiles of X will define the range.  X0.05= –  Z0.95 ≈ 1.2 – (0.02)\*(1.645) ≈ 1.17  X0.95= +  Z0.95 ≈ 1.2 + (0.02)\*(1.645) ≈ 1.23  Where Z0.95 = – Z0.05 = normsinv(1–0.05)=normsinv(0.95)≈1.645 |

2. A local department store wants to investigate the average age of the adults in its existing marketing area to help target its advertising. The store assumes the age follows a normal distribution with a mean of 35 years and a standard deviation of 4 years.

|  |
| --- |
| a. What percentage of adults will be older than 40 years?  P[X>40| =35, =4] = P[(X–)/>(40–35)/4] = P[Z>1.25]≈1–0.89435 ≈ 0.10565  There are 10.565% of the population greater than 40 years old. |
| b. What ages represent the 5th and 95th percentiles?  X0.05= –  Z0.95 ≈ 35 – (4)\*(1.645) ≈ 28.42  X0.95= +  Z0.95 ≈ 35 + (4)\*(1.645) ≈ 41.58 |

|  |  |  |
| --- | --- | --- |
|  | **Common Standard Normal Variates** |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Normal Distribution X~N( ,  2 )  Standard Normal Distribution, Z~N(=0,  2=1)  Standard Normal Variates, Z , where  = P[ Z < Z ]   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | 0.1 | 0.05 | 0.025 | 0.01 | 0.005 | |  | 0.90 | 0.95 | 0.975 | 0.99 | 0.995 | | Z | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 |   . . .   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  | 0.1587 | 0.0668 | 0.0228 | 0.0062 | 0.0013 | |  | 0.8413 | 0.9332 | 0.9772 | 0.9938 | 0.9987 | | Z | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 |   . . . |