NMEC 103 Statistics for Economists 1

May 2021

Formulae list and Statistical tables

1. ANOVA

MEASURES OF CENTRAL TENDENCY

a. Mean
   i. Ungrouped data

$$\boldsymbol{\mu} = \frac{^{x}\mathbf{1} + {}^{x}\mathbf{2} + {}^{x}\mathbf{3} \ldots\ldots {}^{x}\boldsymbol{n}}{\boldsymbol{N}}$$

$$\boldsymbol{\mu} = \frac{\sum_{i=1}^{N} \boldsymbol{x_i}}{\boldsymbol{N}} = \frac{\sum \boldsymbol{x_i}}{\boldsymbol{N}}$$

$$\overline{\boldsymbol{x}} = \frac{\sum \boldsymbol{x_i}}{\boldsymbol{n}}$$

   ii. Grouped data

   Mean $\overline{\boldsymbol{x}} = \frac{\sum \boldsymbol{f_i m_i}}{\sum \boldsymbol{f}}$ $\boldsymbol{or}$ $\frac{\sum \boldsymbol{f_i x_i}}{\boldsymbol{n}}$

   f_i - class frequency

   $m_i \, or \, x_i$ - class mark

   n - number of observations(total frequency)

   $\sum$ - sum of frequency

b. Mode

i. Grouped data

$$\boldsymbol{M_o} = \boldsymbol{O_{mo}} + \frac{\boldsymbol{C}(\boldsymbol{f_m} - \boldsymbol{f_{m-1}}}{\boldsymbol{2f_m} - \boldsymbol{f_{m-1}} - \boldsymbol{f_{m+1}}}$$

$\boldsymbol{O_{mo}}$ - lower limit of the modal class interval

C - class width of the modal class interval

$\boldsymbol{f_m}$ - Frequency of the modal class

$\boldsymbol{f_{m-1}}$- frequency of the interval preceding the modal class/interval

$\boldsymbol{f_{m+1}}$ – frequency of the interval/class following the modal class/interval

c. Median
   i. Ungrouped data

   Position of the median value $\frac{\boldsymbol{n+1}}{\boldsymbol{2}}$th

   ii. Grouped data

- **Grouped data**
- $^{n}/_{2}$to find median interval using cumulative frequency.

$$M_e = o_{me} + \frac{c\left(\frac{n}{2} - f(<)\right)}{f_{me}}$$

o_{me} – lower limit of the median class

C – class width

n – sample size

f_{me} - frequency of the median interval

$f(<)$ - Cumulative frequency of the class before the median class

MEASURES OF POSITION

a. Quartiles

i. Ungrouped data

Q $= \frac{Q(n+1)}{4}$

Q- Quartile of interest

n- number of observations

ii. Grouped Data

$n/4$th position

- $Q_1 = O_{q1} + \frac{c\left(\frac{n}{4} - f(<)\right)}{fq1}$

O_{q1} - lower quartile

n - sample size

f(<) - cumulative frequency of interval before the quartiles

c - class interval

f_q - frequency of the quartile class

b. Percentile

i. Ungrouped data

$$\frac{p(n+1)}{100}$$

p- percentile of interest

n- number of observations

ii. Grouped data

$$\frac{pn}{100} th\ position$$

P = $o_{pi} + \frac{c\left(\frac{n}{100} - f(<)\right)}{f_{pi}}$

o_{pi} - lower limit of the percentile interval

C - interval width

n - sample space

f (<) - cumulative frequency of the interval before p_i

f_{pi} - frequency of the interval class

MEASURES OF DISPERSION

a. Variance

i. Ungrouped data

- Variance $(s^2) = \frac{\left(\sum_{i=1}^{n} x^2 - \frac{(\sum x)^2}{n}\right)}{n-1}$

iii. Grouped data

$$s^2 = \frac{\left(\sum_{i=1}^{n} fx^2 - \frac{(\sum fx)^2}{n}\right)}{n-1}$$

b. Standard Deviation (s)

i. Ungrouped data

$$S = \sqrt{\frac{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)}{n-1}}$$

ii. Grouped data

$$\sqrt{\frac{\left(\sum fx^2 - \frac{(\sum fx)^2}{n}\right)}{n-1}}$$

REGRESSION AND CORRELATION ANALYSIS

Regression equation

$Y = b_0 + b_1X + e$

- Y is the dependent variable,
- X is the independent variable,
- b_0 is the intercept and is constant,
- $b_1$is the slope,

$$b_1 = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

$$b_0 = \frac{\sum Y - b_1 \sum X}{n}$$

Pearson's correlation coefficient (r)

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

Coefficient of determination

$$r^2 = \left(\frac{n\sum XY - \sum X \sum Y}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}\right)^2$$

PROBABILITY

$P(A) = \frac{r}{n}$ where A is an Event, r is number of outcomes of event A, n is total number of all possible outcomes (sample space) and P(A) is the probability of event A occurring.

i. Binomial distribution

$$P(x) = n\,Cx.P^x\,(1-P)^{n-x} = \frac{n!}{x!(n-x)!}P^x(1-P)^{n-x}$$

Where

n is the sample size,

x is number of successes,

P is probability of success and

1-P is the probability of failure

ii. Poisson distribution

$$P(X) = \frac{e^{-a}a^x}{x!}\ where;$$

P(X)is the probablity of success over a given time or space,

a is the expected number of successes per time or space, $a > 0$

e ~ 2.71828 (base of natural logarithms)

iii. Normal Distribution

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

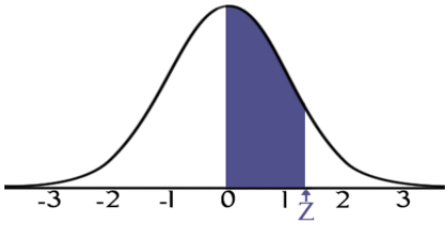
HYPOTHESIS TESTING

i. Mean

$$\frac{\bar{X}-\mu}{\frac{s}{\sqrt{n}}} = Z_c; Sample\ statistic$$

ii. Proportions

Test statistic $Z = \frac{\rho-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$



STANDARD NORMAL TABLE (Z)

Entries in the table give the area under the curve between the mean and *z* standard deviations above the mean. For example, for $z = 1.25$ the area under the curve between the mean (0) and z is 0.3944.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0190	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2969	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3513	0.3554	0.3577	0.3529	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998