STATISTICS HELP CARD

Summary Measures

Sample Mean

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum x_i}{n}$$

Sample Standard Deviation

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

Probability Rules

Complement Rule: $P(A^c) = 1 - P(A)$

Addition Rule: $P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$

Conditional Probability: $P(A|B) = \frac{P(A\ and\ B)}{P(B)}$

Events *A* and *B* are independent if $P(A|B) = P(A)$
Events *A* and *B* are independent if $P(A\ and\ B) = P(A)P(B)$
If *A* and *B* are disjoint events then $P(A\ and\ B) = 0$

General Discrete Random Variable

Mean $E(X) = \mu = \sum x_i p_i = x_1 p_1 + x_2 p_2 + \cdots + x_k p_k$

Standard Deviation $s.d.(X) = \sigma = \sqrt{\sum (x_i - \mu)^2 p_i}$

Standard Score

$$Standard\ Score = \frac{Observation - Mean}{Standard\ Deviation}$$

Z Score

If *X* follows a Normal distribution
with mean μ and standard deviation σ,
then the random variable $Z = \frac{X-\mu}{\sigma}$
has a $N(0,1)$ distribution

Sample Proportion

$$\hat{p} = \frac{x}{n}$$

Mean $E(\hat{p}) = p$

Standard Deviation $s.d.(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$

Sampling Distribution of $\hat{p}$

If the sample size *n* is large enough
(namely, $np \geq 10\ and\ n(1-p) \geq 10$),
then the distribution of all possible sample
proportion values is *approximately*

$$N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

Sample Mean

$$\bar{x} = \frac{\sum x_i}{n}$$

Mean $E(\bar{X}) = \mu$

Standard Deviation $s.d.(\bar{X}) = \frac{\sigma}{\sqrt{n}}$

Sampling Distribution of $\bar{X}$

If *X* has Normal distribution
with mean μ and standard deviation σ,
then the distribution of all possible sample
mean values is

$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Central Limit Theorem

If *X* follows *any* distribution
with mean μ and standard deviation σ
and the sample size *n* is large enough,
then the distribution of all possible sample
mean values is *approximately*

$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

One Population Proportion	Difference in Two Population Proportions	One Population Mean	Population Mean of Differences
Parameter p **Statistic** $\hat{p}$	**Parameter** $p_1 - p_2$ **Statistic** $\hat{p}_1 - \hat{p}_2$	**Parameter** μ **Statistic** $\bar{x}$	**Parameter** μ_d **Statistic** $\bar{d}$
Standard Error $s.e.(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	**Standard Error** $s.e.(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	**Standard Error** $s.e.(\bar{x}) = \frac{s}{\sqrt{n}}$	**Standard Error** $s.e.(\bar{x}_d) = \frac{s_d}{\sqrt{n}}$
Confidence Interval $\hat{p} \pm z^* \times s.e.(\hat{p})$ **Conservative Confidence Interval** $\hat{p} \pm \frac{z^*}{2\sqrt{n}}$ **Sample Size** $n = \left(\frac{z^*}{2m}\right)^2$ *m*=desired margin of error	**Confidence Interval** $(\hat{p}_1 - \hat{p}_2) \pm z^* \times s.e.(\hat{p}_1 - \hat{p}_2)$	**Confidence Interval** $\bar{x} \pm t^* \times s.e.(\bar{x})$ $df = n - 1$	**Confidence Interval** $\bar{x}_d \pm t^* \times s.e.(\bar{x}_d)$ $df = n - 1$
Large Sample z-Test $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$	**Large Sample z-Test** $z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ where $\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$	**One-Sample t-Test** $t = \frac{\bar{x} - \mu_o}{s.e.(\bar{x})}$ $df = n - 1$	**Paired t-Test** $t = \frac{\bar{x}_d - 0}{s.e.(\bar{x}_d)}$ $df = n - 1$

Difference in Two Population Means	
Unpooled (Welch's)	**Pooled**
Parameter $\mu_1 - \mu_2$ **Statistic** $\bar{x}_1 - \bar{x}_2$	**Parameter** $\mu_1 - \mu_2$ **Statistic** $\bar{x}_1 - \bar{x}_2$
Standard Error $s.e.(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	**Standard Error** $pooled\ s.e.(\bar{x}_1 - \bar{x}_2) = s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ where $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$
Confidence Interval $(\bar{x}_1 - \bar{x}_2) \pm t^* \times s.e.(\bar{x}_1 - \bar{x}_2)$ *df* from technology **	**Confidence Interval** $(\bar{x}_1 - \bar{x}_2) \pm t^* \times \left(pooled\ s.e.(\bar{x}_1 - \bar{x}_2)\right)$ $df = n_1 + n_2 - 2$
Two-Sample t-Test $t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{s.e.(\bar{x}_1 - \bar{x}_2)} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ *df* from technology ** ***If technology not available, use conservative df = the minimum of* $n_1 - 1$ *and* $n_2 - 1$	**Pooled Two-Sample t-Test** $t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{pooled\ s.e.(\bar{x}_1 - \bar{x}_2)} = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ $df = n_1 + n_2 - 2$

Note: A z-distribution is often used in statistical methods in place of a t-distribution when sample sizes are sufficiently large.

Pearson Correlation and Linear Regression

Pearson Correlation and its square

$$r = \sum \left(\frac{x - \bar{x}}{s_x}\right)\left(\frac{y - \bar{y}}{s_y}\right)$$

$r^2 = \frac{SSReg}{SSTotal}$ where $SSTotal = \sum(y - \bar{y})^2 = SSReg + SSE$

Estimate of σ

$$s = \sqrt{MSE} = \sqrt{\frac{SSE}{n-2}}$$

where $SSE = \sum(y - \hat{y})^2 = \sum e^2$

Linear Regression Model

Population Version

Mean: $E(Y|x) = \beta_0 + \beta_1 x$

Individual: $y_i = \beta_0 + \beta_1 x + \varepsilon_i$

where ε_i is $N(0, \sigma)$

Sample Version

Mean: $\hat{y} = b_0 + b_1 x$

Individual: $y_i = b_0 + b_1 x + e_i$

Standard Error of the Sample Slope

$$s.e.(b_1) = \frac{s}{\sqrt{\sum(x - \bar{x})^2}}$$

Confidence Interval for $\boldsymbol{\beta_1}$

$b_1 \pm t^* \times s.e.(b_1)$ *df* = $n - 2$

t-Test for $\boldsymbol{\beta_1}$

$t = \frac{b_1 - 0}{s.e.(b_1)}$ *df* = $n - 2$

Parameter Estimators

$$b_1 = r\frac{s_y}{s_x}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

Confidence Interval for the Mean Response

$\hat{y} \pm t^* \times s.e.(fit)$ *df* = $n - 2$

where $s.e.(fit) = s\sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum(x_i - \bar{x})^2}}$

Residuals

$$e = y - \hat{y} = observed\ y - predicted\ y$$

Prediction Interval for an Individual Response

$\hat{y} \pm t^* \times s.e.(pred)$ *df* = $n - 2$

where $s.e.(pred) = \sqrt{s^2 + \left(s.e.(fit)\right)^2}$

Chi-Square Tests

Test for Goodness of Fit

Expected Count

$$Expected = np_{i0}$$

Test Statistic

$\chi^2 = \sum \frac{(observed - expected)^2}{expected}$ $df = k - 1$

Test of Independence

Expected Count

$$Expected = \frac{(row\ total)(column\ total)}{total\ n}$$

Test Statistic

$\chi^2 = \sum \frac{(observed - expected)^2}{expected}$ $df = (r - 1)(c - 1)$

Properties of a Chi-Square Distribution

A χ^2 random variable has mean = *df* and standard deviation = $\sqrt{2df}$

STATISTICS HELP CARD TABLES

Z Table: Table entry for z is area to the left of z

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

T Table: t^* multiplers for confidence intervals and bounds for tail probabilities

df	0.50	0.75	0.80	0.85	0.90	0.95	0.98	0.99	0.998	0.999
1	1.000	2.414	3.078	4.165	6.314	12.706	31.821	63.657	318.309	636.619
2	0.816	1.604	1.886	2.282	2.920	4.303	6.965	9.925	22.327	31.599
3	0.765	1.423	1.638	1.924	2.353	3.182	4.541	5.841	10.215	12.924
4	0.741	1.344	1.533	1.778	2.132	2.776	3.747	4.604	7.173	8.610
5	0.727	1.301	1.476	1.699	2.015	2.571	3.365	4.032	5.893	6.869
6	0.718	1.273	1.440	1.650	1.943	2.447	3.143	3.707	5.208	5.959
7	0.711	1.254	1.415	1.617	1.895	2.365	2.998	3.499	4.785	5.408
8	0.706	1.240	1.397	1.592	1.860	2.306	2.896	3.355	4.501	5.041
9	0.703	1.230	1.383	1.574	1.833	2.262	2.821	3.250	4.297	4.781
10	0.700	1.221	1.372	1.559	1.812	2.228	2.764	3.169	4.144	4.587
11	0.697	1.214	1.363	1.548	1.796	2.201	2.718	3.106	4.025	4.437
12	0.695	1.209	1.356	1.538	1.782	2.179	2.681	3.055	3.930	4.318
13	0.694	1.204	1.350	1.530	1.771	2.160	2.650	3.012	3.852	4.221
14	0.692	1.200	1.345	1.523	1.761	2.145	2.624	2.977	3.787	4.140
15	0.691	1.197	1.341	1.517	1.753	2.131	2.602	2.947	3.733	4.073
16	0.690	1.194	1.337	1.512	1.746	2.120	2.583	2.921	3.686	4.015
17	0.689	1.191	1.333	1.508	1.740	2.110	2.567	2.898	3.646	3.965
18	0.688	1.189	1.330	1.504	1.734	2.101	2.552	2.878	3.610	3.922
19	0.688	1.187	1.328	1.500	1.729	2.093	2.539	2.861	3.579	3.883
20	0.687	1.185	1.325	1.497	1.725	2.086	2.528	2.845	3.552	3.850
21	0.686	1.183	1.323	1.494	1.721	2.080	2.518	2.831	3.527	3.819
22	0.686	1.182	1.321	1.492	1.717	2.074	2.508	2.819	3.505	3.792
23	0.685	1.180	1.319	1.489	1.714	2.069	2.500	2.807	3.485	3.768
24	0.685	1.179	1.318	1.487	1.711	2.064	2.492	2.797	3.467	3.745
25	0.684	1.178	1.316	1.485	1.708	2.060	2.485	2.787	3.450	3.725
26	0.684	1.177	1.315	1.483	1.706	2.056	2.479	2.779	3.435	3.707
27	0.684	1.176	1.314	1.482	1.703	2.052	2.473	2.771	3.421	3.690
28	0.683	1.175	1.313	1.480	1.701	2.048	2.467	2.763	3.408	3.674
29	0.683	1.174	1.311	1.479	1.699	2.045	2.462	2.756	3.396	3.659
30	0.683	1.173	1.310	1.477	1.697	2.042	2.457	2.750	3.385	3.646
31	0.682	1.172	1.309	1.476	1.696	2.040	2.453	2.744	3.375	3.633
32	0.682	1.172	1.309	1.475	1.694	2.037	2.449	2.738	3.365	3.622
33	0.682	1.171	1.308	1.474	1.692	2.035	2.445	2.733	3.356	3.611
34	0.682	1.170	1.307	1.473	1.691	2.032	2.441	2.728	3.348	3.601
35	0.682	1.170	1.306	1.472	1.690	2.030	2.438	2.724	3.340	3.591
36	0.681	1.169	1.306	1.471	1.688	2.028	2.434	2.719	3.333	3.582
37	0.681	1.169	1.305	1.470	1.687	2.026	2.431	2.715	3.326	3.574
38	0.681	1.168	1.304	1.469	1.686	2.024	2.429	2.712	3.319	3.566
39	0.681	1.168	1.304	1.468	1.685	2.023	2.426	2.708	3.313	3.558
40	0.681	1.167	1.303	1.468	1.684	2.021	2.423	2.704	3.307	3.551
50	0.679	1.164	1.299	1.462	1.676	2.009	2.403	2.678	3.261	3.496
60	0.679	1.162	1.296	1.458	1.671	2.000	2.390	2.660	3.232	3.460
70	0.678	1.160	1.294	1.456	1.667	1.994	2.381	2.648	3.211	3.435
80	0.678	1.159	1.292	1.453	1.664	1.990	2.374	2.639	3.195	3.416
90	0.677	1.158	1.291	1.452	1.662	1.987	2.368	2.632	3.183	3.402
100	0.677	1.157	1.290	1.451	1.660	1.984	2.364	2.626	3.174	3.390
500	0.675	1.152	1.283	1.442	1.648	1.965	2.334	2.586	3.107	3.310
1000	0.675	1.151	1.282	1.441	1.646	1.962	2.330	2.581	3.098	3.300
z*	0.674	1.150	1.282	1.440	1.645	1.960	2.326	2.576	3.090	3.291
Tail prob	**0.25**	**0.125**	**0.10**	**0.075**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**

Chi-Square Table: Table entry is χ^2 value with area to the right equal to column heading

df	0.50	0.30	0.25	0.20	0.10	0.075	0.05	0.025	0.01	0.005	0.001
1	0.455	1.074	1.323	1.642	2.706	3.170	3.841	5.024	6.635	7.879	10.828
2	1.386	2.408	2.773	3.219	4.605	5.181	5.991	7.378	9.210	10.597	13.816
3	2.366	3.665	4.108	4.642	6.251	6.905	7.815	9.348	11.345	12.838	16.266
4	3.357	4.878	5.385	5.989	7.779	8.496	9.488	11.143	13.277	14.860	18.467
5	4.351	6.064	6.626	7.289	9.236	10.008	11.070	12.833	15.086	16.750	20.515
6	5.348	7.231	7.841	8.558	10.645	11.466	12.592	14.449	16.812	18.548	22.458
7	6.346	8.383	9.037	9.803	12.017	12.883	14.067	16.013	18.475	20.278	24.322
8	7.344	9.524	10.219	11.030	13.362	14.270	15.507	17.535	20.090	21.955	26.124
9	8.343	10.656	11.389	12.242	14.684	15.631	16.919	19.023	21.666	23.589	27.877
10	9.342	11.781	12.549	13.442	15.987	16.971	18.307	20.483	23.209	25.188	29.588
11	10.341	12.899	13.701	14.631	17.275	18.294	19.675	21.920	24.725	26.757	31.264
12	11.340	14.011	14.845	15.812	18.549	19.602	21.026	23.337	26.217	28.300	32.909
13	12.340	15.119	15.984	16.985	19.812	20.897	22.362	24.736	27.688	29.819	34.528
14	13.339	16.222	17.117	18.151	21.064	22.180	23.685	26.119	29.141	31.319	36.123
15	14.339	17.322	18.245	19.311	22.307	23.452	24.996	27.488	30.578	32.801	37.697
16	15.338	18.418	19.369	20.465	23.542	24.716	26.296	28.845	32.000	34.267	39.252
17	16.338	19.511	20.489	21.615	24.769	25.970	27.587	30.191	33.409	35.718	40.790
18	17.338	20.601	21.605	22.760	25.989	27.218	28.869	31.526	34.805	37.156	42.312
19	18.338	21.689	22.718	23.900	27.204	28.458	30.144	32.852	36.191	38.582	43.820
20	19.337	22.775	23.828	25.038	28.412	29.692	31.410	34.170	37.566	39.997	45.315
21	20.337	23.858	24.935	26.171	29.615	30.920	32.671	35.479	38.932	41.401	46.797
22	21.337	24.939	26.039	27.301	30.813	32.142	33.924	36.781	40.289	42.796	48.268
23	22.337	26.018	27.141	28.429	32.007	33.360	35.172	38.076	41.638	44.181	49.728
24	23.337	27.096	28.241	29.553	33.196	34.572	36.415	39.364	42.980	45.559	51.179
25	24.337	28.172	29.339	30.675	34.382	35.780	37.652	40.646	44.314	46.928	52.620
26	25.336	29.246	30.435	31.795	35.563	36.984	38.885	41.923	45.642	48.290	54.052
27	26.336	30.319	31.528	32.912	36.741	38.184	40.113	43.195	46.963	49.645	55.476
28	27.336	31.391	32.620	34.027	37.916	39.380	41.337	44.461	48.278	50.993	56.892
29	28.336	32.461	33.711	35.139	39.087	40.573	42.557	45.722	49.588	52.336	58.301
30	29.336	33.530	34.800	36.250	40.256	41.762	43.773	46.979	50.892	53.672	59.703