Estimating the True Variance

- The true variance, σ^2, estimated with P % confidence, is in the range

$$\frac{\nu S_x^2}{\chi^2_{(\alpha/2)}} \le \sigma^2 \le \frac{\nu S_x^2}{\chi^2_{1-(\alpha/2)}}$$

noting α = 1 – P and ν = N -1.

In-Class Example (x’ and σ Inference)

- Given the mean and standard deviation are 10 and 1.5, respectively, for a sample of 16, estimate with 95 % confidence the ranges within which are the true mean and true standard deviation.

Using the χ^2 Table

ν	$\chi^2_{0.99}$	$\chi^2_{0.975}$	$\chi^2_{0.95}$	$\chi^2_{0.90}$	$\chi^2_{0.50}$	$\chi^2_{0.05}$	$\chi^2_{0.025}$	$\chi^2_{0.01}$
1	0.000	0.000	0.000	0.016	0.455	3.84	5.02	6.63
2	0.020	0.051	0.103	0.211	1.39	5.99	7.38	9.21
3	0.115	0.216	0.352	0.584	2.37	7.81	9.35	11.3
4	0.297	0.484	0.711	1.06	3.36	9.49	11.1	13.3
5	0.554	0.831	1.15	1.61	4.35	11.1	12.8	15.1
6	0.872	1.24	1.64	2.20	5.35	12.6	14.4	16.8
7	1.24	1.69	2.17	2.83	6.35	14.1	16.0	18.5
8	1.65	2.18	2.73	3.49	7.34	15.5	17.5	20.1
9	2.09	2.70	3.33	4.17	8.34	16.9	19.0	21.7
10	2.56	3.25	3.94	4.78	9.34	18.3	20.5	23.2
11	3.05	3.82	4.57	5.58	10.3	19.7	21.9	24.7
12	3.57	4.40	5.23	6.30	11.3	21.0	23.3	26.2
13	4.11	5.01	5.89	7.04	12.3	22.4	24.7	27.7
14	4.66	5.63	6.57	7.79	13.3	23.7	26.1	29.1
15	5.23	6.26	7.26	8.55	14.3	25.0	27.5	30.6
16	5.81	6.91	7.96	9.31	15.3	26.3	28.8	32.0
17	6.41	7.56	8.67	10.1	16.3	27.6	30.2	33.4
18	7.01	8.23	9.39	10.9	17.3	28.9	31.5	34.8
19	7.63	8.91	10.1	11.7	18.3	30.1	32.9	36.2
20	8.26	9.59	10.9	12.4	19.3	31.4	34.2	37.6
30	15.0	16.8	18.5	20.6	29.3	43.8	47.0	50.9
40	22.2	24.4	26.5	29.1	39.3	55.8	59.3	63.7
50	29.7	32.4	34.8	37.7	49.3	67.5	71.4	76.2
60	37.5	40.5	43.2	46.5	59.3	79.1	83.3	88.4
70	45.4	48.8	51.7	55.3	69.3	90.5	95.0	100.4
80	53.5	57.2	60.4	64.3	79.3	101.9	106.6	112.3
90	61.8	65.6	69.1	73.3	89.3	113.1	118.1	124.1
100	70.1	74.2	77.9	82.4	99.3	124.3	129.6	135.8

Establishing a Rejection Criterion

- There is a probability α (= 1 – P), for a sample of size N with sample variance S_x^2 drawn from a population with true variance σ^2, that the difference between S_x^2 and σ^2 is solely due to random effects.

$$\chi^2 \equiv \nu S_x^2 / \sigma^2 = \chi_\alpha^2$$

- For example, there is only a 5 % probability that a value of χ^2 = 25.0 would result solely due to random effects for a sample of N = 16 (>> ν = 15), as found from the χ^2 table.

In-Class Example (Rejection Criterion)

- The sample standard deviation of the length of 12 widgets taken off of an assembly lines is 0.20 mm. What must be the widget population's standard deviation to support the conclusion that the probability is 50 % for any difference between the sample's and population's standard deviations to be the result of random effects ?

Comparing a Sample and Population

using *chinormchk.m*

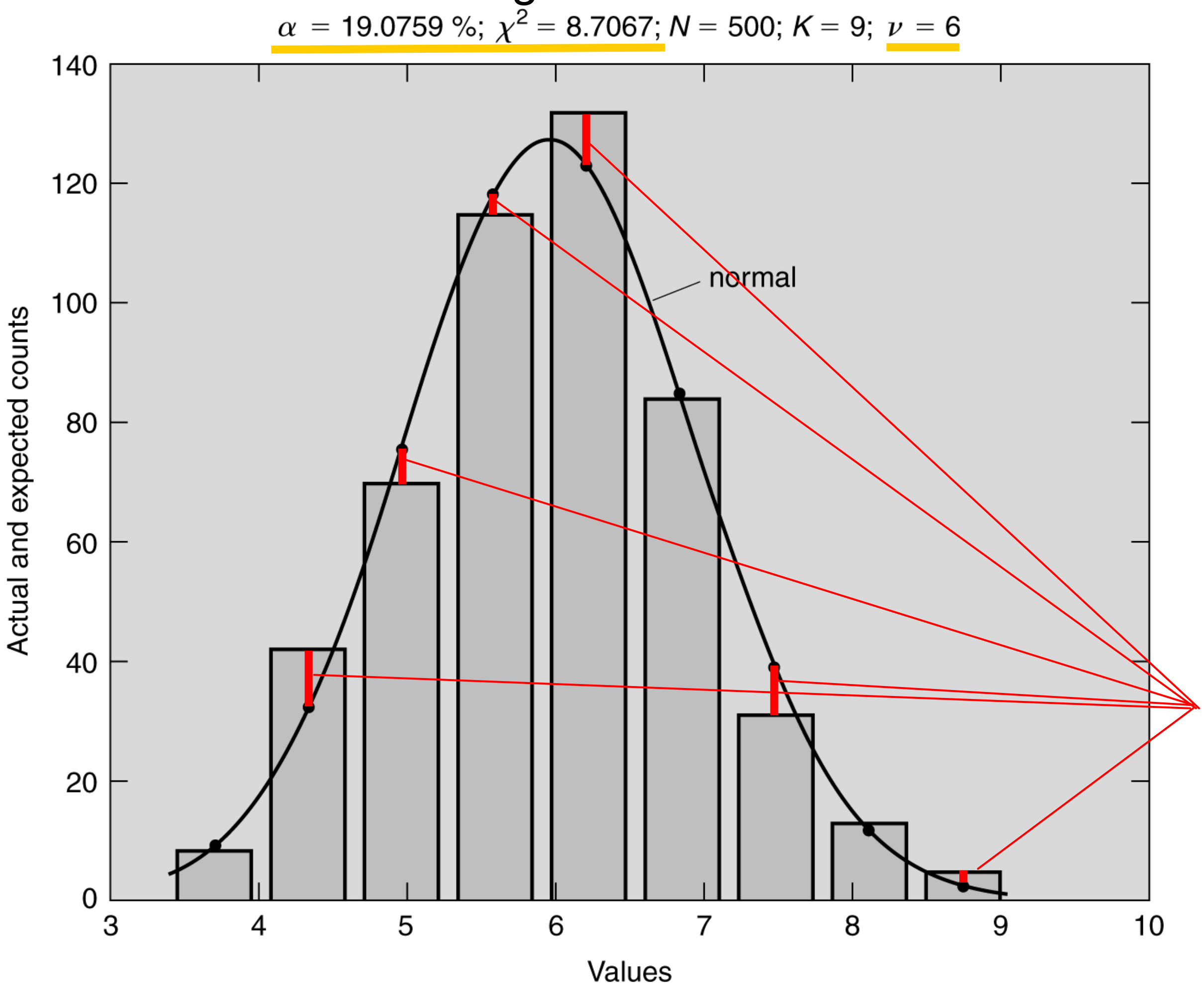

Figure 8.13

Procedure

K (=1.15N$^{1/3}$) bins

cover range ≥ ~±2σ

$$E_j = NP_j$$

$$\chi^2 \approx \sum_{j=1}^{K} \frac{(O_j - E_j)^2}{E_j}$$

→ α (for ν = K-3)

Example Problem 8.10

Statement: Consider a study conducted by a professor who wishes to determine whether or not the 300 undergraduate engineering students in his department are normal. He determines this by comparing the distribution of their heights to that expected for a normally distributed student population. His height data are presented in Table 8.9.

Solution: For this case, $\nu = 8 - 3 = 5$, where $K = 8$ was determined using Scott's formula (actually, $K = 7.7$, which then is rounded up). The expected values are calculated for each bin by noting that $E_k = NP_k$ where $N = 300$. For example, for bin 2 where $-1.5\sigma \leq x \leq -\sigma$, $P_k = P(-1.5\sigma \leq x \leq -\sigma) = P(z_1 = -1.5) - P(z_1 = -1) = 0.4332 - 0.3413$ (using Table 8.2) $= 0.0919$. So, the expected number in bin 2 is $(0.0919)(300) = 27.5$. The results for every bin are shown in Table 8.9.

Substitution of these results into Equation 8.46 yields $\chi^2 = 0.904$. For the values of $\chi^2_\alpha =$

0.904 and $\nu = 5$, from Table 8.8, $\alpha \simeq 0.97$. That is, the probability of obtaining this χ^2 value or less is ~97 %, under the assumption that the expected distribution is correct. Thus, agreement with the assumed normal distribution is *significant*. This is contrary to what the professor originally thought!

Bin number k	Heights in bin	Observed number, O_k	Expected number, E_k
1	less than $X - 1.5\sigma$	19	20.1
2	between $X - 1.5\sigma$ and $X - \sigma$	25	27.5
3	between $X - \sigma$ and $X - 0.5\sigma$	44	45.0
4	between $X - 0.5\sigma$ and X	59	57.5
5	between X and $X + 0.5\sigma$	60	57.5
6	between $X + 0.5\sigma$ and $X + \sigma$	45	45.0
7	between $X + \sigma$ and $X + 1.5\sigma$	30	27.5
8	above $X + 1.5\sigma$	18	20.1

Table 8.9: Observed and expected heights