Nonparametric ANOVA; Homogeneity of Variances

I. **Nonparametric ANOVA – Kruskal-Wallis test** (≅ ANOVA by ranks)

A. Application: • whenever parametric single-factor ANOVA appropriate,
• when parametric parametric assumptions violated: non-normal populations; ≠ variances

B. Assumption: sampled populations have same dispersions & shapes

C. Hypotheses: H_0: measurements the same in all k populations
H_A: measurements <u>not</u> the same in all k populations

D. Procedure: Rank data – same as with Mann-Whitney two-sample test

E. Test statistic:

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} - 3(N+1)$$

n_i = # in group i; R_i=sum ranks in i

F. Compare w/ critical values, Table B.13

G. If k > 5 or n_i large (> 6 or 7), approximate critical values of H: X^2 w/ k-1 DF

H. If tied ranks, H an underestimate. Correction factor:

$$C = 1 - \frac{\sum t}{N^3 - N}$$

H_c = H/C

$$\sum t = \sum_{i=1}^{m}(t_i^3 - t_i)$$

t_i = # ties in i^{th} group, m= # groups of tied ranks

$H_c \approx H$ when t_i's << N

II. **Test for Homogeneity of Variances – Bartlett's test**

A. Hypotheses:

H_0: $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$

H_A: variances are not all equal

B. Test Statistic:

$$B = (\ln s_p^2)\left(\sum_{i=1}^{k}\nu_i\right) - \sum_{i=1}^{k}\nu_i \ln s_i^2$$

$\nu_i = n_i - 1$, s_p^2 = pooled variance = $\sum SS_i / \sum \nu_i$

C. B ≈ Chi-squared; critical value: $\chi^2_{\alpha,k-1}$, where k = number of samples

E. More accurate w/ correction factor:

$$C = 1 + \frac{1}{3(k-1)}\left(\sum\frac{1}{\nu_i} - \frac{1}{\sum\nu_i}\right)$$

Corrected statistic: $B_c = B/C$ Compare B_c with $\chi^2_{\alpha,k-1}$

F. Test badly affected by deviations from normality
Not recommended to test assumptions prior to doing single-factor ANOVA,
– esp. since ANOVA relatively robust to unequal variances.