HW7 solution

1) Generate 500 random numbers between 0 and 1, based on the linear congruential techniques covered last week. ***Print no more than 50 of them***. Use a modulus size of at least 500. Find the number of up-runs and down-runs in the sequence. Using the run-test method discussed, see if the sequence generated appears independent with a 10% level of significance.

$X_0 := 23 \qquad a := 7 \qquad c := 15 \qquad m := 607$

$N := 500$

$i := 1 .. N$

$X_i := \mathrm{mod}\left[\left(X_{i-1} \cdot a + c\right), m\right]$

upruns(r) :=
- for $i \in 0 .. \mathrm{length}(r) - 2$
  - $x_i \leftarrow \mathrm{if}\left(r_{i+1} > r_i, 1, 0\right)$
- $k \leftarrow 1$
- $ups \leftarrow \mathrm{if}\left(x_0 = 1, 1, 0\right)$
- while $(k < \mathrm{length}(x))$
  - $ups \leftarrow \mathrm{if}\left[\left(x_{k-1} = 0\right) \wedge \left(x_k = 1\right), ups + 1, ups\right]$
  - $k \leftarrow k + 1$
- return ups

$\mathrm{downruns}(x) := \mathrm{upruns}(-x)$

$\mathrm{upruns}(X) = 143 \qquad \mathrm{downruns}(X) = 143$

$\mathrm{runs}(x) := \mathrm{upruns}(x) + \mathrm{downruns}(x)$

$\mathrm{runs}(X) = 286$

$\mu_a := \frac{2 \cdot N - 1}{3} \qquad \mu_a = 333$

$\sigma_a := \sqrt{\frac{16 \cdot N - 29}{90}} \qquad \sigma_a = 9.411$

$Z_0 := \frac{\mathrm{runs}(X) - \mu_a}{\sigma_a} \qquad Z_0 = -4.994$

For 10% level of significance,

$$\text{qnorm}(.05, 0, 1) = -1.645$$

Or, calculated alternatively:

The cumulative normal distribution is:

$$\phi(z_a) := \int_{-\infty}^{z_a} \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{\frac{-u^2}{2}} \, du$$

Find a z.a so that ϕ(-z.a) = .05 and ϕ(z.a)=.95. This will provide the critical value z.a that we can compare the run statistic z.0 to.

$$\phi(-1.645) = 0.05$$

$$\phi(1.645) = 0.95$$

so, $z_{0.5} := 1.645$

The calculated statistic is -04.994, which is not in the range of -z.α - z.α, so we can reject the null hypothesis. This distribution of random numbers can be distinguished from a uniform distribution at the 10% level of significance.

2) Use the built-in random number generator in whatever programming environment you are familiar with (C, Matlab, Excel, Mathcad, etc.). Generate as many random numbers as you can in a reasonable period of time (e.g., 5-10 minutes of computer execution time) and apply at least one of the tests for randomness we have discussed (***limit your homework submission for this problem to 2-3 pages at most. I don't want to have to print 50 pages to give to the grader***). Extra credit for more than one randomness test.

$N := 10000000$

$X := \text{runif}(N, 0, 1)$

RUNS TEST:

$\text{runs}(X) = 6.666 \times 10^6$

$\mu_a := \frac{2 \cdot N - 1}{3} \qquad \mu_a = 6.667 \times 10^6$

$\sigma_a := \sqrt{\frac{16 \cdot N - 29}{90}} \qquad \sigma_a = 1.333 \times 10^3$

$Z_0 := \frac{\text{runs}(X) - \mu_a}{\sigma_a} \qquad Z_0 = -0.671$

For 5% level of significance,

$$\text{qnorm}(.025, 0, 1) = -1.96$$

Or, calculated alternatively:

The cumulative normal distribution is:

$$\phi(z_a) := \int_{-\infty}^{z_a} \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{\frac{-u^2}{2}} \, du$$

Find a z.a so that ϕ(-z.a) = .025 and ϕ(z.a)=.975. This will provide the critical value z.a that we can compare the run statistic z.0 to.

$$\phi(-1.96) = 0.025$$

$$\phi(1.96) = 0.975$$

so, $z_{0.5} := 1.96$

The calculated statistic is in the range of -z.α - z.α, so we cannot reject the null hypothesis. This distribution of random numbers cannot be distinguished from a uniform distribution at the 5% level of significance.

CHI-SQUARED TEST:

$k := 50 \quad p := \frac{1}{k} \qquad p = 0.02$

$i := 0 .. k - 1 \qquad a_i := \frac{i}{k}$

$a_k := 1 \qquad E_i := p \cdot N$

$O := hist(a, X) \qquad term_i := \frac{(O_i - E_i)^2}{E_i}$

$$\chi_{sq} := \sum_{i=0}^{k-1} \frac{(O_i - E_i)^2}{E_i} \qquad \chi_{sq} = 54.334$$

$\chi_{sq.critical.5\%} := 67.5$

Since the calculated chi-squared is less than the critical value, the null hypothesis cannot be rejected.

3) Extra credit: If you have access to a different type of computing environment, try the same experiment with the same method and compare your results. For instance, if you use Excel, you could use the same spreadsheet on a PC vs. a Macintosh vs. StarOffice with linux. If you use Matlab or C, you could try the same program on Sun/UNIX, SGI/Irix, PC/Windows, PC/linux. Different x86 PCs with different versions of Windows doesn't count.