**CALIFORNIA STATE UNIVERSITY, HAYWARD**
**DEPARTMENT OF STATISTICS**

**Statistics 3601 Introductory Statistics for Scientists and Engineers**
**Winter 2001**

**Computer Exam #2** Due Wednesday, February 28

Suppose we have a population of interest whose distribution is UNKOWN, but its expected value, $\mu$ , and its standard deviation, $\sigma$ , are both KNOWN. The Central Limit Theorem says, if we repeatedly take samples of size *n* from that population, regardless of its shape, then the sampling distribution of

$$Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

will be approximately Normal(0,1) and the approximation improves as *n* increases.

To see the CLT in action simulate repeated independent sampling from the following distributions:

1. $Normal(\mu = 10, \sigma^2 = 9)$
2. $Exponential(\theta = 10)$
3. $Uniform(a = 0, b = 20)$

Simulate 2000 independent random samples from each distribution for each f the following samples sizes *n* = 10, 30, 50.

To check the approximate Normality of the sampling distribution of the $\overline{X}$ 's, when $\mu$ and $\sigma$ are both KNOWN, compare the proportion of *Z*'s that fall into the following intervals with the true portions calculated form the Standard Normal distribution.

(-∞,-3], (-3, -2], (-2, -1], (-1, 0], (0, 1], (1, 2], (2, 3], (3, ∞)

(The true proportions are given below.) Comment on the accuracy of the CLT for the different sample sizes from each population.

True proportions:

| | (-∞,-3] | (-3, -2] | (-2, -1] | (-1, 0] | (0, 1] | (1, 2] | (2, 3] | (3, ∞) |
|---|---|---|---|---|---|---|---|---|
| *Z* | 0.0013 | 0.0215 | 0.1359 | 0.3413 | 0.3413 | 0.1359 | 0.0215 | 0.0013 |