```
------------------------------------------------------------------------------
      name:  <unnamed>
       log:  C:\Users\saiomkark\OneDrive - The University of Chicago\AdvStats\
> PS5\Question3.log
  log type:  text
 opened on:   5 Nov 2021, 23:18:58

. *
. * clear memory in stata
. *
. clear

. *
. *Use Stata in this exercise. Suppose that x is drawn from the following "mix
> ing distribution." Let y be a binary random variable with Pr(y = 1) = 0.9. I
> f y = 1, then x is drawn from a standard normal distribution. If y = 0, then
>  x is drawn from a normal distribution with μ = 100 and standard deviation σ
>  = 20.
. * a) Find the mean of x.
. * b) For this distribution, use 10,000 draws from each of the following samp
> le sizes: n = 36, n = 64, n = 100, n = 225, n = 2500, and n = 12100. Discuss
>  how well the normal approximation fits your simulated estimates of the mean
>  at the critical values of 0.025 and 0.975.
.
.
. set seed 10072021

. set obs 100000
number of observations (_N) was 0, now 100,000

. * y is binary with Pr(y1=1) = 0.9
. *
. gen y = runiform() < 0.9

. sum y

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
           y |    100,000      .90021    .2997213          0          1

. *
. * If y = 1, then x is drawn from a standard normal distribution. If y = 0, t
> hen x is drawn from a normal distribution with μ = 100 and standard deviatio
> n σ = 20
. *
. gen x = rnormal() if (y == 1)
(9,979 missing values generated)

. replace x = rnormal(100, 20) if (y == 0)
(9,979 real changes made)

. sum x

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
           x |    100,000     10.0146    30.76765  -4.810123   171.9138

.
. clear
```

```
.
. *
. * In case the program "clt" is resident in memory drop it because we will re
> define it in this program.  Capture simply tells stata to ignore this comman
> d if the "clt" is not in memory
. *
. capture program drop clt

. *
. * define program.  This is the program that we will 10,000 times for each sa
> mple size; rclass tells us that we will be using results stored in the class
>  r()
. *
. program clt, rclass
  1. version 15.0
  2. *
. * Simulations for central limit theorem (clt)
. *
. * declare argument of program; only one in this program the sample size
. *
. args N
  3. *
. * clear any variables in memory
. *
. clear
  4. *
. * set observations to `N'. "quietly" tells stata not to put the results on y
> our screen
. *
. quietly set obs `N'
  5. *
. * Set temporary variables.  Temporary variables are dropped whenever we star
> t a new "loop" through the data
. *
. tempvar y x
  6. *
. * y is binary with Pr(y1=1) = 0.9
. *
. gen `y' = uniform() < 0.9
  7. *
. * If y = 1, then x is drawn from a standard normal distribution. If y = 0, t
> hen x is drawn from a normal distribution with μ = 100 and standard deviatio
> n σ = 20
. *
. gen `x' = rnormal() if (`y' == 1)
  8. replace `x' = rnormal(100, 20) if (`y' == 0)
  9.
. *
. * calculate means and standard deviation; "quietly" command suppresses outpu
> t
. *
. quietly sum `y'
 10. *
. * return scalar mu"y" recovers the mean sd the standard deviation
. *
. return scalar muy = r(mean)
 11. return scalar sdy = r(sd)
 12. *
. * calculate means and standard deviation; "quietly" command suppresses outpu
> t
```

```
. *
. quietly sum `x'
 13. *
. * return scalar mu"x" recovers the mean sd the standard deviation
. *
. return scalar mux = r(mean)
 14. return scalar sdx = r(sd)
 15. *
. * end the program definition
. *
. end

. *
. * Set out seed
. *
. set seed 24031997

.
. *
. * run simulation for 36 observations
. *
. simulate muy = r(muy) mux = r(mux) sdy = r(sdy) sdx = r(sdx), reps(10000) no
> dots: clt 36

      command:  clt 36
          muy:  r(muy)
          mux:  r(mux)
          sdy:  r(sdy)
          sdx:  r(sdx)


. *
. * save this data
. *
. save clt36, replace
file clt36.dta saved

. *
. * Sample using 36 observations
. *
. use clt36
(simulate: clt)

. gen z1 = (muy-0.9)      /(sdy/36^(1/2))
(221 missing values generated)

. gen z2 = (mux-10)      /(sdx/36^(1/2))

.
. *
. * construct rejection level 5 percent rate; left and right tails
. *
. gen rrej1 = z1 >  1.96 & z1 < .

. gen lrej1 = z1 < -1.96

. gen rrej2 = z2 >  1.96

. gen lrej2 = z2 < -1.96
```

```
. *
. * Should be 0.05, 1 binary 2 binary
. *
. sum rrej1-lrej2 // 36 observations

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .0876    .2827264          0          1
       lrej1 |     10,000       .0081    .0896392          0          1
       rrej2 |     10,000        .006    .0772308          0          1
       lrej2 |     10,000       .1031    .3041046          0          1

. /*
> 99% (-1%)  2.33
> 95% (-5%)  1.65
> 90% (-10%) 1.29
> 75% (-25%) 0.58
> */
. bitest lrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000         81          250       0.02500      0.00810

  Pr(k >= 81)             = 1.000000  (one-sided test)
  Pr(k <= 81)             = 0.000000  (one-sided test)
  Pr(k <= 81 or k >= 469) = 0.000000  (two-sided test)

. bitest rrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000        876          250       0.02500      0.08760

  Pr(k >= 876) = 0.000000  (one-sided test)
  Pr(k <= 876) = 1.000000  (one-sided test)
  Pr(k >= 876) = 0.000000  (two-sided test)

  note: lower tail of two-sided p-value is empty

. bitest lrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       1031          250       0.02500      0.10310

  Pr(k >= 1031) = 0.000000  (one-sided test)
  Pr(k <= 1031) = 1.000000  (one-sided test)
  Pr(k >= 1031) = 0.000000  (two-sided test)

  note: lower tail of two-sided p-value is empty

. bitest rrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000         60          250       0.02500      0.00600

  Pr(k >= 60)             = 1.000000  (one-sided test)
  Pr(k <= 60)             = 0.000000  (one-sided test)
  Pr(k <= 60 or k >= 508) = 0.000000  (two-sided test)
```

```
.
. replace rrej1 = z1 >  0.975 & z1 < .
(1,778 real changes made)

. replace lrej1 = z1 < -0.975
(1,393 real changes made)

. replace rrej2 = z2 >  0.975
(1,147 real changes made)

. replace lrej2 = z2 < -0.975
(1,345 real changes made)

.
. *
. * Should be around 0.175 given that ~35% of sample means remain outliers at
> z = 1 (~ 17.5% on each side) for a normal distribution
. *
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .2654     .441568          0          1
       lrej1 |     10,000       .1474    .3545219          0          1
       rrej2 |     10,000       .1207    .3257946          0          1
       lrej2 |     10,000       .2376    .4256341          0          1

.
. bitest lrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       1474         1750       0.17500      0.14740

  Pr(k >= 1474)              = 1.000000  (one-sided test)
  Pr(k <= 1474)              = 0.000000  (one-sided test)
  Pr(k <= 1474 or k >= 2038) = 0.000000  (two-sided test)

. bitest rrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       2654         1750       0.17500      0.26540

  Pr(k >= 2654)             = 0.000000  (one-sided test)
  Pr(k <= 2654)             = 1.000000  (one-sided test)
  Pr(k <= 955 or k >= 2654) = 0.000000  (two-sided test)

. bitest lrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       2376         1750       0.17500      0.23760

  Pr(k >= 2376)              = 0.000000  (one-sided test)
  Pr(k <= 2376)              = 1.000000  (one-sided test)
  Pr(k <= 1177 or k >= 2376) = 0.000000  (two-sided test)

. bitest rrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       1207         1750       0.17500      0.12070

  Pr(k >= 1207)              = 1.000000  (one-sided test)
  Pr(k <= 1207)              = 0.000000  (one-sided test)
  Pr(k <= 1207 or k >= 2342) = 0.000000  (two-sided test)
```

```
.
.
. replace rrej1 = z1 >  0.025 & z1 < .
(2,266 real changes made)

. replace lrej1 = z1 < -0.025
(3,385 real changes made)

. replace rrej2 = z2 >  0.025
(3,333 real changes made)

. replace lrej2 = z2 < -0.025
(2,913 real changes made)

.
. *
. * Should be 0.50
. *
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000        .492     .499961          0          1
       lrej1 |     10,000       .4859    .4998261          0          1
       rrej2 |     10,000        .454    .4979044          0          1
       lrej2 |     10,000       .5289    .4991891          0          1

. *
. * Should be around 0.50 on each side given that value of z is low at 0.025.
> Most sample means should fall as outliers for this range.
. *
. bitest lrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       4859         5000       0.50000      0.48590

  Pr(k >= 4859)              = 0.997674  (one-sided test)
  Pr(k <= 4859)              = 0.002476  (one-sided test)
  Pr(k <= 4859 or k >= 5141) = 0.004952  (two-sided test)

. bitest rrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       4920         5000       0.50000      0.49200

  Pr(k >= 4920)              = 0.946303  (one-sided test)
  Pr(k <= 4920)              = 0.055915  (one-sided test)
  Pr(k <= 4920 or k >= 5080) = 0.111830  (two-sided test)

. bitest lrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       5289         5000       0.50000      0.52890

  Pr(k >= 5289)              = 0.000000  (one-sided test)
  Pr(k <= 5289)              = 1.000000  (one-sided test)
  Pr(k <= 4711 or k >= 5289) = 0.000000  (two-sided test)
```

```
. bitest rrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       4540         5000       0.50000      0.45400

  Pr(k >= 4540)              = 1.000000  (one-sided test)
  Pr(k <= 4540)              = 0.000000  (one-sided test)
  Pr(k <= 4540 or k >= 5460) = 0.000000  (two-sided test)

.
. clear

.
. *
. *  run simulation for 64 observations
. *
. simulate muy = r(muy) mux = r(mux) sdy = r(sdy) sdx = r(sdx), reps(10000) no
> dots: clt 64

      command:  clt 64
          muy:  r(muy)
          mux:  r(mux)
          sdy:  r(sdy)
          sdx:  r(sdx)


. *
. * save this data
. *
. save clt64, replace
file clt64.dta saved

. *
. * Sample using 64 observations
. *
. use clt64
(simulate: clt)

. gen z1 = (muy-0.9)      /(sdy/64^(1/2))
(11 missing values generated)

. gen z2 = (mux-10)      /(sdx/64^(1/2))

.
. *
. * construct rejection level 5 percent rate; left and right tails
. *
. gen rrej1 = z1 >  1.96 & z1 < .

. gen lrej1 = z1 < -1.96

. gen rrej2 = z2 >  1.96

. gen lrej2 = z2 < -1.96

. *
. * Should be 0.05, 1 binary 2 binary
. *
. sum rrej1-lrej2 // 64 observations

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .1048    .3063109          0          1
       lrej1 |     10,000       .0096     .097513          0          1
       rrej2 |     10,000       .0082    .0901863          0          1
       lrej2 |     10,000        .076    .2650114          0          1
```

```
. /*
> 99% (-1%)  2.33
> 95% (-5%)  1.65
> 90% (-10%) 1.29
> 75% (-25%) 0.58
> */
. bitest lrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000         96          250       0.02500      0.00960

  Pr(k >= 96)             = 1.000000  (one-sided test)
  Pr(k <= 96)             = 0.000000  (one-sided test)
  Pr(k <= 96 or k >= 444) = 0.000000  (two-sided test)

. bitest rrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       1048          250       0.02500      0.10480

  Pr(k >= 1048) = 0.000000  (one-sided test)
  Pr(k <= 1048) = 1.000000  (one-sided test)
  Pr(k >= 1048) = 0.000000  (two-sided test)

  note: lower tail of two-sided p-value is empty

. bitest lrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000        760          250       0.02500      0.07600

  Pr(k >= 760) = 0.000000  (one-sided test)
  Pr(k <= 760) = 1.000000  (one-sided test)
  Pr(k >= 760) = 0.000000  (two-sided test)

  note: lower tail of two-sided p-value is empty

. bitest rrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000         82          250       0.02500      0.00820

  Pr(k >= 82)             = 1.000000  (one-sided test)
  Pr(k <= 82)             = 0.000000  (one-sided test)
  Pr(k <= 82 or k >= 467) = 0.000000  (two-sided test)

.
. replace rrej1 = z1 >  0.975 & z1 < .
(1,131 real changes made)

. replace lrej1 = z1 < -0.975
(970 real changes made)

. replace rrej2 = z2 >  0.975
(1,208 real changes made)
```

```
. replace lrej2 = z2 < -0.975
(1,360 real changes made)

.
. *
. * Should be around 0.175 given that ~35% of sample means remain outliers at
> z = 1 (~ 17.5% on each side) for a normal distribution
. *
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .2179    .4128397          0          1
       lrej1 |     10,000       .1066    .3086194          0          1
       rrej2 |     10,000        .129    .3352167          0          1
       lrej2 |     10,000        .212    .4087453          0          1

.
. bitest lrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       1066         1750       0.17500      0.10660

  Pr(k >= 1066)             = 1.000000  (one-sided test)
  Pr(k <= 1066)             = 0.000000  (one-sided test)
  Pr(k <= 1066 or k >= 2514) = 0.000000  (two-sided test)

. bitest rrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       2179         1750       0.17500      0.21790

  Pr(k >= 2179)             = 0.000000  (one-sided test)
  Pr(k <= 2179)             = 1.000000  (one-sided test)
  Pr(k <= 1346 or k >= 2179) = 0.000000  (two-sided test)

. bitest lrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       2120         1750       0.17500      0.21200

  Pr(k >= 2120)             = 0.000000  (one-sided test)
  Pr(k <= 2120)             = 1.000000  (one-sided test)
  Pr(k <= 1398 or k >= 2120) = 0.000000  (two-sided test)

. bitest rrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       1290         1750       0.17500      0.12900

  Pr(k >= 1290)             = 1.000000  (one-sided test)
  Pr(k <= 1290)             = 0.000000  (one-sided test)
  Pr(k <= 1290 or k >= 2244) = 0.000000  (two-sided test)

.
.
```

```
. replace rrej1 = z1 >  0.025 & z1 < .
(3,223 real changes made)

. replace lrej1 = z1 < -0.025
(3,521 real changes made)

. replace rrej2 = z2 >  0.025
(3,292 real changes made)

. replace lrej2 = z2 < -0.025
(3,087 real changes made)

.
. *
. * Should be 0.50
. *
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .5402    .4984063          0          1
       lrej1 |     10,000       .4587    .4983163          0          1
       rrej2 |     10,000       .4582    .4982746          0          1
       lrej2 |     10,000       .5207    .4995963          0          1

.
. *
. * Should be around 0.50 on each side given that value of z is low at 0.025.
> Most sample means should fall as outliers for this range.
. *
.
. bitest lrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       4587         5000       0.50000      0.45870

  Pr(k >= 4587)              = 1.000000  (one-sided test)
  Pr(k <= 4587)              = 0.000000  (one-sided test)
  Pr(k <= 4587 or k >= 5413) = 0.000000  (two-sided test)

. bitest rrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       5402         5000       0.50000      0.54020

  Pr(k >= 5402)              = 0.000000  (one-sided test)
  Pr(k <= 5402)              = 1.000000  (one-sided test)
  Pr(k <= 4598 or k >= 5402) = 0.000000  (two-sided test)

. bitest lrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       5207         5000       0.50000      0.52070

  Pr(k >= 5207)              = 0.000018  (one-sided test)
  Pr(k <= 5207)              = 0.999983  (one-sided test)
  Pr(k <= 4793 or k >= 5207) = 0.000036  (two-sided test)
```

```
. bitest rrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       4582         5000       0.50000      0.45820

  Pr(k >= 4582)              = 1.000000  (one-sided test)
  Pr(k <= 4582)              = 0.000000  (one-sided test)
  Pr(k <= 4582 or k >= 5418) = 0.000000  (two-sided test)

.
. clear

. *
. *  run simulation for 100 observations
. *
. simulate muy = r(muy) mux = r(mux) sdy = r(sdy) sdx = r(sdx), reps(10000) no
> dots: clt 100

      command:  clt 100
          muy:  r(muy)
          mux:  r(mux)
          sdy:  r(sdy)
          sdx:  r(sdx)


. *
. * save this data
. *
. save clt100, replace
file clt100.dta saved

. *
. * Sample using 100 observations
. *
. use clt100
(simulate: clt)

. gen z1 = (muy-0.9)      /(sdy/100^(1/2))
(1 missing value generated)

. gen z2 = (mux-10)      /(sdx/100^(1/2))

. *
. * construct rejection level 5 percent rate; left and right tails
. *
. gen rrej1 = z1 >  1.96 & z1 < .

. gen lrej1 = z1 < -1.96

. gen rrej2 = z2 >  1.96

. gen lrej2 = z2 < -1.96

. *
. * Should be 0.05, 1 binary 2 binary
. *
. sum rrej1-lrej2 // 100 observations

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .0536    .2252379          0          1
       lrej1 |     10,000       .0101    .0999949          0          1
       rrej2 |     10,000       .0096     .097513          0          1
       lrej2 |     10,000       .0582     .234133          0          1
```

```
. /*
> 99% (-1%)  2.33
> 95% (-5%)  1.65
> 90% (-10%) 1.29
> 75% (-25%) 0.58
> */
. bitest lrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000        101          250       0.02500      0.01010

  Pr(k >= 101)             = 1.000000  (one-sided test)
  Pr(k <= 101)             = 0.000000  (one-sided test)
  Pr(k <= 101 or k >= 436) = 0.000000  (two-sided test)

. bitest rrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000        536          250       0.02500      0.05360

  Pr(k >= 536)            = 0.000000  (one-sided test)
  Pr(k <= 536)            = 1.000000  (one-sided test)
  Pr(k <= 46 or k >= 536) = 0.000000  (two-sided test)

. bitest lrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000        582          250       0.02500      0.05820

  Pr(k >= 582)            = 0.000000  (one-sided test)
  Pr(k <= 582)            = 1.000000  (one-sided test)
  Pr(k <= 26 or k >= 582) = 0.000000  (two-sided test)

. bitest rrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000         96          250       0.02500      0.00960

  Pr(k >= 96)             = 1.000000  (one-sided test)
  Pr(k <= 96)             = 0.000000  (one-sided test)
  Pr(k <= 96 or k >= 444) = 0.000000  (two-sided test)

.
. replace rrej1 = z1 >  0.975 & z1 < .
(1,469 real changes made)

. replace lrej1 = z1 < -0.975
(1,189 real changes made)

. replace rrej2 = z2 >  0.975
(1,283 real changes made)

. replace lrej2 = z2 < -0.975
(1,378 real changes made)

.
```

```
. *
. * Should be around 0.175 given that ~35% of sample means remain outliers at
> z = 1 (~ 17.5% on each side) for a normal distribution
. *
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .2005    .4003945          0          1
       lrej1 |     10,000        .129    .3352167          0          1
       rrej2 |     10,000       .1379    .3448122          0          1
       lrej2 |     10,000        .196    .3969884          0          1

.
. bitest lrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       1290         1750       0.17500      0.12900

  Pr(k >= 1290)              = 1.000000  (one-sided test)
  Pr(k <= 1290)              = 0.000000  (one-sided test)
  Pr(k <= 1290 or k >= 2244) = 0.000000  (two-sided test)

. bitest rrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       2005         1750       0.17500      0.20050

  Pr(k >= 2005)              = 0.000000  (one-sided test)
  Pr(k <= 2005)              = 1.000000  (one-sided test)
  Pr(k <= 1503 or k >= 2005) = 0.000000  (two-sided test)

. bitest lrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       1960         1750       0.17500      0.19600

  Pr(k >= 1960)              = 0.000000  (one-sided test)
  Pr(k <= 1960)              = 1.000000  (one-sided test)
  Pr(k <= 1545 or k >= 1960) = 0.000000  (two-sided test)

. bitest rrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       1379         1750       0.17500      0.13790

  Pr(k >= 1379)              = 1.000000  (one-sided test)
  Pr(k <= 1379)              = 0.000000  (one-sided test)
  Pr(k <= 1379 or k >= 2143) = 0.000000  (two-sided test)

.
. *
. * Should be around 0.50 on each side given that value of z is low at 0.025.
> Most sample means should fall as outliers for this range.
. *
.
```

```
. replace rrej1 = z1 >  0.025 & z1 < .
(2,488 real changes made)

. replace lrej1 = z1 < -0.025
(2,846 real changes made)

. replace rrej2 = z2 >  0.025
(3,306 real changes made)

. replace lrej2 = z2 < -0.025
(3,130 real changes made)

.
. *
. * Should be 0.50
. *
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .4493    .4974477          0          1
       lrej1 |     10,000       .4136    .4925031          0          1
       rrej2 |     10,000       .4685    .4990317          0          1
       lrej2 |     10,000        .509     .499944          0          1

.
. bitest lrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       4136         5000       0.50000      0.41360

  Pr(k >= 4136)              = 1.000000  (one-sided test)
  Pr(k <= 4136)              = 0.000000  (one-sided test)
  Pr(k <= 4136 or k >= 5864) = 0.000000  (two-sided test)

. bitest rrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       4493         5000       0.50000      0.44930

  Pr(k >= 4493)              = 1.000000  (one-sided test)
  Pr(k <= 4493)              = 0.000000  (one-sided test)
  Pr(k <= 4493 or k >= 5507) = 0.000000  (two-sided test)

. bitest lrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       5090         5000       0.50000      0.50900

  Pr(k >= 5090)              = 0.036724  (one-sided test)
  Pr(k <= 5090)              = 0.964855  (one-sided test)
  Pr(k <= 4910 or k >= 5090) = 0.073449  (two-sided test)

. bitest rrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       4685         5000       0.50000      0.46850

  Pr(k >= 4685)              = 1.000000  (one-sided test)
  Pr(k <= 4685)              = 0.000000  (one-sided test)
  Pr(k <= 4685 or k >= 5315) = 0.000000  (two-sided test)
```

```
.
. clear

.
. *
. *  run simulation for 225 observations
. *
. simulate muy = r(muy) mux = r(mux) sdy = r(sdy) sdx = r(sdx), reps(10000) no
> dots: clt 225

      command:  clt 225
          muy:  r(muy)
          mux:  r(mux)
          sdy:  r(sdy)
          sdx:  r(sdx)


. *
. * save this data
. *
. save clt225, replace
file clt225.dta saved

. *
. * Sample using 225 observations
. *
. use clt225
(simulate: clt)

. gen z1 = (muy-0.9)      /(sdy/225^(1/2))

. gen z2 = (mux-10)      /(sdx/225^(1/2))

.
. *
. * construct rejection level 5 percent rate; left and right tails
. *
. gen rrej1 = z1 >  1.96 & z1 < .

. gen lrej1 = z1 < -1.96

. gen rrej2 = z2 >  1.96

. gen lrej2 = z2 < -1.96

. *
. * Should be 0.05, 1 binary 2 binary
. *
. sum rrej1-lrej2 // 225 observations

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .0515    .2210263          0          1
       lrej1 |     10,000       .0179    .1325947          0          1
       rrej2 |     10,000       .0142    .1183206          0          1
       lrej2 |     10,000       .0439    .2048829          0          1

. /*
> 99% (-1%)  2.33
> 95% (-5%)  1.65
> 90% (-10%) 1.29
> 75% (-25%) 0.58
> */
```

```
. bitest lrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000        179          250       0.02500      0.01790

  Pr(k >= 179)             = 0.999999  (one-sided test)
  Pr(k <= 179)             = 0.000001  (one-sided test)
  Pr(k <= 179 or k >= 328) = 0.000002  (two-sided test)

. bitest rrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000        515          250       0.02500      0.05150

  Pr(k >= 515)            = 0.000000  (one-sided test)
  Pr(k <= 515)            = 1.000000  (one-sided test)
  Pr(k <= 56 or k >= 515) = 0.000000  (two-sided test)

. bitest lrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000        439          250       0.02500      0.04390

  Pr(k >= 439)            = 0.000000  (one-sided test)
  Pr(k <= 439)            = 1.000000  (one-sided test)
  Pr(k <= 98 or k >= 439) = 0.000000  (two-sided test)

. bitest rrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000        142          250       0.02500      0.01420

  Pr(k >= 142)             = 1.000000  (one-sided test)
  Pr(k <= 142)             = 0.000000  (one-sided test)
  Pr(k <= 142 or k >= 375) = 0.000000  (two-sided test)

.
. replace rrej1 = z1 >  0.975 & z1 < .
(1,345 real changes made)

. replace lrej1 = z1 < -0.975
(1,107 real changes made)

. replace rrej2 = z2 >  0.975
(1,208 real changes made)

. replace lrej2 = z2 < -0.975
(1,459 real changes made)

. *
. * Should be around 0.175 given that ~35% of sample means remain outliers at
> z = 1 (~ 17.5% on each side) for a normal distribution
. *
.
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000        .186    .3891261          0          1
       lrej1 |     10,000       .1286    .3347734          0          1
       rrej2 |     10,000        .135    .3417407          0          1
       lrej2 |     10,000       .1898    .3921624          0          1
```

```
.
. bitest lrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       1286         1750       0.17500      0.12860

  Pr(k >= 1286)              = 1.000000  (one-sided test)
  Pr(k <= 1286)              = 0.000000  (one-sided test)
  Pr(k <= 1286 or k >= 2249) = 0.000000  (two-sided test)

. bitest rrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       1860         1750       0.17500      0.18600

  Pr(k >= 1860)              = 0.002107  (one-sided test)
  Pr(k <= 1860)              = 0.998059  (one-sided test)
  Pr(k <= 1641 or k >= 1860) = 0.004118  (two-sided test)

. bitest lrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       1898         1750       0.17500      0.18980

  Pr(k >= 1898)              = 0.000061  (one-sided test)
  Pr(k <= 1898)              = 0.999945  (one-sided test)
  Pr(k <= 1604 or k >= 1898) = 0.000115  (two-sided test)

. bitest rrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       1350         1750       0.17500      0.13500

  Pr(k >= 1350)              = 1.000000  (one-sided test)
  Pr(k <= 1350)              = 0.000000  (one-sided test)
  Pr(k <= 1350 or k >= 2176) = 0.000000  (two-sided test)

.
.
. replace rrej1 = z1 >  0.025 & z1 < .
(3,259 real changes made)

. replace lrej1 = z1 < -0.025
(3,595 real changes made)

. replace rrej2 = z2 >  0.025
(3,415 real changes made)

. replace lrej2 = z2 < -0.025
(3,145 real changes made)

.
. *
. * Should be around 0.50 on each side given that value of z is low at 0.025.
> Most sample means should fall as outliers for this range.
. *
```

```
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .5119    .4998834          0          1
       lrej1 |     10,000       .4881    .4998834          0          1
       rrej2 |     10,000       .4765    .4994724          0          1
       lrej2 |     10,000       .5043    .5000065          0          1

.
. bitest lrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       4881         5000       0.50000      0.48810

  Pr(k >= 4881)              = 0.991578  (one-sided test)
  Pr(k <= 4881)              = 0.008892  (one-sided test)
  Pr(k <= 4881 or k >= 5119) = 0.017784  (two-sided test)

. bitest rrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       5119         5000       0.50000      0.51190

  Pr(k >= 5119)              = 0.008892  (one-sided test)
  Pr(k <= 5119)              = 0.991578  (one-sided test)
  Pr(k <= 4881 or k >= 5119) = 0.017784  (two-sided test)

. bitest lrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       5043         5000       0.50000      0.50430

  Pr(k >= 5043)              = 0.197663  (one-sided test)
  Pr(k <= 5043)              = 0.807849  (one-sided test)
  Pr(k <= 4957 or k >= 5043) = 0.395326  (two-sided test)

. bitest rrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       4765         5000       0.50000      0.47650

  Pr(k >= 4765)              = 0.999999  (one-sided test)
  Pr(k <= 4765)              = 0.000001  (one-sided test)
  Pr(k <= 4765 or k >= 5235) = 0.000003  (two-sided test)

.
. clear

.
. *
. *  run simulation for 2500 observations
. *
. simulate muy = r(muy) mux = r(mux) sdy = r(sdy) sdx = r(sdx), reps(10000) no
> dots: clt 2500

      command:  clt 2500
          muy:  r(muy)
          mux:  r(mux)
          sdy:  r(sdy)
          sdx:  r(sdx)
```

```
. *
. * save this data
. *
. save clt2500, replace
file clt2500.dta saved

. *
. * Sample using 2500 observations
. *
. use clt2500
(simulate: clt)

. gen z1 = (muy-0.9)      /(sdy/2500^(1/2))

. gen z2 = (mux-10)      /(sdx/2500^(1/2))

.
. *
. * construct rejection level 5 percent rate; left and right tails
. *
. gen rrej1 = z1 >  1.96 & z1 < .

. gen lrej1 = z1 < -1.96

. gen rrej2 = z2 >  1.96

. gen lrej2 = z2 < -1.96

. *
. * Should be 0.05, 1 binary 2 binary
. *
. sum rrej1-lrej2 // 2500 observations

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .0309    .1730554          0          1
       lrej1 |     10,000       .0229    .1495922          0          1
       rrej2 |     10,000        .021    .1433913          0          1
       lrej2 |     10,000       .0301     .170871          0          1

. /*
> 99% (-1%)  2.33
> 95% (-5%)  1.65
> 90% (-10%) 1.29
> 75% (-25%) 0.58
> */
. bitest lrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000        229          250       0.02500      0.02290

  Pr(k >= 229)             = 0.917242  (one-sided test)
  Pr(k <= 229)             = 0.093273  (one-sided test)
  Pr(k <= 229 or k >= 271) = 0.189024  (two-sided test)

. bitest rrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000        309          250       0.02500      0.03090

  Pr(k >= 309)             = 0.000143  (one-sided test)
  Pr(k <= 309)             = 0.999886  (one-sided test)
  Pr(k <= 194 or k >= 309) = 0.000257  (two-sided test)
```

```
. bitest lrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000        301          250      0.02500      0.03010

  Pr(k >= 301)            = 0.000826  (one-sided test)
  Pr(k <= 301)            = 0.999329  (one-sided test)
  Pr(k <= 201 or k >= 301) = 0.001507  (two-sided test)

. bitest rrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000        210          250      0.02500      0.02100

  Pr(k >= 210)            = 0.996058  (one-sided test)
  Pr(k <= 210)            = 0.004810  (one-sided test)
  Pr(k <= 210 or k >= 292) = 0.009459  (two-sided test)

.
. replace rrej1 = z1 >  0.975 & z1 < .
(1,364 real changes made)

. replace lrej1 = z1 < -0.975
(1,284 real changes made)

. replace rrej2 = z2 >  0.975
(1,390 real changes made)

. replace lrej2 = z2 < -0.975
(1,434 real changes made)

.
. *
. * Should be around 0.175 given that ~35% of sample means remain outliers at
> z = 1 (~ 17.5% on each side) for a normal distribution
. *
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .1673    .3732622          0          1
       lrej1 |     10,000       .1513     .358359          0          1
       rrej2 |     10,000         .16    .3666244          0          1
       lrej2 |     10,000       .1735    .3786979          0          1

.
. bitest lrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       1513         1750      0.17500      0.15130

  Pr(k >= 1513)             = 1.000000  (one-sided test)
  Pr(k <= 1513)             = 0.000000  (one-sided test)
  Pr(k <= 1513 or k >= 1996) = 0.000000  (two-sided test)

. bitest rrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       1673         1750      0.17500      0.16730

  Pr(k >= 1673)             = 0.979763  (one-sided test)
  Pr(k <= 1673)             = 0.021576  (one-sided test)
  Pr(k <= 1673 or k >= 1828) = 0.042711  (two-sided test)
```

```
. bitest lrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       1735         1750       0.17500      0.17350

  Pr(k >= 1735)              = 0.657463  (one-sided test)
  Pr(k <= 1735)              = 0.352280  (one-sided test)
  Pr(k <= 1735 or k >= 1765) = 0.702754  (two-sided test)

. bitest rrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       1600         1750       0.17500      0.16000

  Pr(k >= 1600)              = 0.999969  (one-sided test)
  Pr(k <= 1600)              = 0.000035  (one-sided test)
  Pr(k <= 1600 or k >= 1903) = 0.000070  (two-sided test)

.
.
. replace rrej1 = z1 >  0.025 & z1 < .
(3,259 real changes made)

. replace lrej1 = z1 < -0.025
(3,308 real changes made)

. replace rrej2 = z2 >  0.025
(3,257 real changes made)

. replace lrej2 = z2 < -0.025
(3,203 real changes made)

.
. *
. * Should be around 0.50 on each side given that value of z is low at 0.025.
> Most sample means should fall as outliers for this range.
. *
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .4932    .4999788          0          1
       lrej1 |     10,000       .4821    .4997045          0          1
       rrej2 |     10,000       .4857    .4998205          0          1
       lrej2 |     10,000       .4938    .4999866          0          1

.
. bitest lrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       4821         5000       0.50000      0.48210

  Pr(k >= 4821)              = 0.999835  (one-sided test)
  Pr(k <= 4821)              = 0.000178  (one-sided test)
  Pr(k <= 4821 or k >= 5179) = 0.000357  (two-sided test)

. bitest rrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       4932         5000       0.50000      0.49320

  Pr(k >= 4932)              = 0.914658  (one-sided test)
  Pr(k <= 4932)              = 0.088507  (one-sided test)
  Pr(k <= 4932 or k >= 5068) = 0.177013  (two-sided test)
```

```
. bitest lrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       4938         5000       0.50000      0.49380

  Pr(k >= 4938)              = 0.894351  (one-sided test)
  Pr(k <= 4938)              = 0.109348  (one-sided test)
  Pr(k <= 4938 or k >= 5062) = 0.218695  (two-sided test)

. bitest rrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       4857         5000       0.50000      0.48570

  Pr(k >= 4857)              = 0.997949  (one-sided test)
  Pr(k <= 4857)              = 0.002185  (one-sided test)
  Pr(k <= 4857 or k >= 5143) = 0.004370  (two-sided test)

.
. clear

.
. *
. *  run simulation for 12100 observations
. *
. simulate muy = r(muy) mux = r(mux) sdy = r(sdy) sdx = r(sdx), reps(10000) no
> dots: clt 12100

      command:  clt 12100
          muy:  r(muy)
          mux:  r(mux)
          sdy:  r(sdy)
          sdx:  r(sdx)


. *
. * save this data
. *
. save clt12100, replace
file clt12100.dta saved

. *
. *
. * Sample using 12100 observations
. *
. use clt12100
(simulate: clt)

. gen z1 = (muy-0.9)      /(sdy/12100^(1/2))

. gen z2 = (mux-10)      /(sdx/12100^(1/2))

.
. *
. * construct rejection level 5 percent rate; left and right tails
. *
. gen rrej1 = z1 >  1.96 & z1 < .
```

```
. gen lrej1 = z1 < -1.96

. gen rrej2 = z2 >  1.96

. gen lrej2 = z2 < -1.96

. *
. * Should be 0.05, 1 binary 2 binary
. *
. sum rrej1-lrej2 // 12100 observations

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .0265    .1606248          0          1
       lrej1 |     10,000       .0236    .1518069          0          1
       rrej2 |     10,000       .0236    .1518069          0          1
       lrej2 |     10,000       .0278    .1644076          0          1

. /*
> 99% (-1%)  2.33
> 95% (-5%)  1.65
> 90% (-10%) 1.29
> 75% (-25%) 0.58
> */
. bitest lrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000        236          250       0.02500      0.02360

  Pr(k >= 236)            = 0.823156  (one-sided test)
  Pr(k <= 236)            = 0.194287  (one-sided test)
  Pr(k <= 236 or k >= 264) = 0.387168  (two-sided test)

. bitest rrej1 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000        265          250       0.02500      0.02650

  Pr(k >= 265)            = 0.176122  (one-sided test)
  Pr(k <= 265)            = 0.839667  (one-sided test)
  Pr(k <= 234 or k >= 265) = 0.336526  (two-sided test)

. bitest lrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000        278          250       0.02500      0.02780

  Pr(k >= 278)            = 0.040818  (one-sided test)
  Pr(k <= 278)            = 0.964317  (one-sided test)
  Pr(k <= 222 or k >= 278) = 0.078024  (two-sided test)

. bitest rrej2 = 0.025

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000        236          250       0.02500      0.02360

  Pr(k >= 236)            = 0.823156  (one-sided test)
  Pr(k <= 236)            = 0.194287  (one-sided test)
  Pr(k <= 236 or k >= 264) = 0.387168  (two-sided test)
```

```
.
. replace rrej1 = z1 >  0.975 & z1 < .
(1,416 real changes made)

. replace lrej1 = z1 < -0.975
(1,416 real changes made)

. replace rrej2 = z2 >  0.975
(1,419 real changes made)

. replace lrej2 = z2 < -0.975
(1,358 real changes made)

.
. *
. * Should be around 0.175 given that ~35% of sample means remain outliers at
> z = 1 (~ 17.5% on each side) for a normal distribution
. *
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .1681    .3739738          0          1
       lrej1 |     10,000       .1652    .3713795          0          1
       rrej2 |     10,000       .1655    .3716498          0          1
       lrej2 |     10,000       .1636    .3699307          0          1

.
. bitest lrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       1652         1750       0.17500      0.16520

  Pr(k >= 1652)             = 0.995461  (one-sided test)
  Pr(k <= 1652)             = 0.004906  (one-sided test)
  Pr(k <= 1652 or k >= 1849) = 0.009897  (two-sided test)

. bitest rrej1 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       1681         1750       0.17500      0.16810

  Pr(k >= 1681)             = 0.966817  (one-sided test)
  Pr(k <= 1681)             = 0.035199  (one-sided test)
  Pr(k <= 1681 or k >= 1820) = 0.069384  (two-sided test)

. bitest lrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       1636         1750       0.17500      0.16360

  Pr(k >= 1636)             = 0.998806  (one-sided test)
  Pr(k <= 1636)             = 0.001304  (one-sided test)
  Pr(k <= 1636 or k >= 1866) = 0.002581  (two-sided test)

. bitest rrej2 = 0.175

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       1655         1750       0.17500      0.16550

  Pr(k >= 1655)             = 0.994280  (one-sided test)
  Pr(k <= 1655)             = 0.006170  (one-sided test)
  Pr(k <= 1655 or k >= 1846) = 0.012404  (two-sided test)
```

```
.
.
. replace rrej1 = z1 >  0.025 & z1 < .
(3,210 real changes made)

. replace lrej1 = z1 < -0.025
(3,337 real changes made)

. replace rrej2 = z2 >  0.025
(3,294 real changes made)

. replace lrej2 = z2 < -0.025
(3,218 real changes made)

.
. *
. * Should be around 0.50 on each side given that value of z is low at 0.025.
> Most sample means should fall as outliers for this range.
. *
. sum rrej1-lrej2

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
       rrej1 |     10,000       .4891    .4999062          0          1
       lrej1 |     10,000       .4989    .5000238          0          1
       rrej2 |     10,000       .4949     .499999          0          1
       lrej2 |     10,000       .4854    .4998118          0          1

.
. bitest lrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej1 |    10000       4989         5000       0.50000      0.49890

  Pr(k >= 4989)              = 0.590953  (one-sided test)
  Pr(k <= 4989)              = 0.416834  (one-sided test)
  Pr(k <= 4989 or k >= 5011) = 0.833669  (two-sided test)

. bitest rrej1 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej1 |    10000       4891         5000       0.50000      0.48910

  Pr(k >= 4891)              = 0.985740  (one-sided test)
  Pr(k <= 4891)              = 0.015001  (one-sided test)
  Pr(k <= 4891 or k >= 5109) = 0.030002  (two-sided test)

. bitest lrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       lrej2 |    10000       4854         5000       0.50000      0.48540

  Pr(k >= 4854)              = 0.998306  (one-sided test)
  Pr(k <= 4854)              = 0.001806  (one-sided test)
  Pr(k <= 4854 or k >= 5146) = 0.003612  (two-sided test)

. bitest rrej2 = 0.50

    Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
       rrej2 |    10000       4949         5000       0.50000      0.49490

  Pr(k >= 4949)              = 0.848495  (one-sided test)
  Pr(k <= 4949)              = 0.156248  (one-sided test)
  Pr(k <= 4949 or k >= 5051) = 0.312495  (two-sided test)
```

```
.
. clear

.
.
. *Observations
. *Central Limit theorem is interpreted here.
. *a)Mean of the random variable 'x' following the 'Mixed distribution' = 10
. *b)
. *1. As we increase N, the percentage of sample means that have a z-score bel
> ow -0.025 and above 0.025 is ~99%.
. *2. For critical point z = 0.975: As we increase N, the percentage of sample
>  means that have a z-score below -0.975 and above 0.975 is ~34%, which means
>  66% of the sample means are between z score of 0.975.
. *These simulation results are in accordance with a typical normal distributi
> on where almost 68% of sample means lie within a z-score of 1 and where many
>  sample means fall outside the z-score of 0.025 as the interval defined by t
> he same is very very small.
.
.
.
. *
. * close log file
. *
. log close
      name:  <unnamed>
       log:  C:\Users\saiomkark\OneDrive - The University of Chicago\AdvStats\
> PS5\Question3.log
  log type:  text
 closed on:   5 Nov 2021, 23:20:02
-------------------------------------------------------------------------------
```