



NAME: _____

1 Objective: Calculating exact coverage probabilities of \hat{p} and \tilde{p} in binomial distribution

Lecture on exact coverage derivation of formulas in class.

Ordinary sample proportion for point estimation: $\hat{p} = S/n$; Adjusted Wald approximation for point estimation $\tilde{p} = (S + 2)/(n + 4)$; S is binomial $S \sim \text{Bin}(n, p)$

2 Key Words

probit, probbnml, coverage probability.

3 Exact coverage probability

The code below computes the exact coverage probability:

```
DATA covprob;
  p = .5;
  q = 1-p;

  z = abs(probit(.025)); /* z_star for 95% */
  zz = z*z;

  do n = 10 to 50;
    w = zz / (2 * n) ;
    rad = sqrt( 2*w*p*q + w*w);
    rt = (p + w + rad)/(1+2*w);
    lt = (p + w - rad)/(1+2*w);

    nhat_rt = floor(n * rt);
    nhat_lt = ceil(n * lt - 1);

    ncurl_rt = floor((n+4)*rt - 2);
    ncurl_lt = ceil((n+4)*lt - 3);

    /* Calculate left cumulative binomial probs */
    p1 = probbnml(p, n, nhat_lt);          /* P[S <= lt] */
    p2 = probbnml(p, n, nhat_rt);
    cv_hat = p2 - p1;

    p3 = probbnml(p, n, ncurl_lt);
    p4 = probbnml(p, n, ncurl_rt);
    cv_curl = p4 - p3;
```

```
        output;
    end;
    keep n cv_hat cv_curl;
run;

PROC print data = covprob;
run;

PROC plot data = covprob;
    plot cv_hat*n = '^';
    plot cv_curl*n = '~';
run;
```

4 Turn In

Write a short paragraph in which you answer the following questions including coverage probability graphs for \hat{p} and \tilde{p} .

1. Compare your results with Lab 04 simulation for $n = 12, 17, 18$.
2. Get better graphs using PROC sgplot. Note the reference line 95%.

```
proc sgplot data=covprob;
title h=2 'Figure_2:_COVERAGE_PROBABILITY';
    series X=n Y=cv_hat;
    reffline .95 /
    axis = y
    label = "95% confidence line"
    lineattrs = ( color = red ) ;
run;
quit;

proc sgplot data=covprob;
title h=2 'Figure_2:_COVERAGE_PROBABILITY';
    series X=n Y=cv_curl;
    reffline .95 /
    axis = y
    label = "95% confidence line"
    lineattrs = ( color = red ) ;
run;
quit;
```

3. Superimpose the the graphs to get a better comparison between `cv_hat` and `cv_curl`.
4. Turn in superimposed graph. Which is better, exact coverage or approximate coverage? Why? Give reasons.