STAT 30040 Homework 2, Exercises 1 & 4

Robert Winter

Table of Contents

1 Exercise 1: Binomial Probability Approximations

1.1 Introduction

Suppose $X \sim \mathrm{Bin}(n, p)$. We are interested in the probability $\mathrm{P}(X = k)$ for:

- **Case 1:** $k = 3$, $n = 6$, $p = 0.3$;

- **Case 2:** $k = 11$, $n = 40$, $p = 0.4$; **and**

- **Case 3:** $k = 2$, $n = 400$, $p = 0.003$;

as listed in the table below. In each case, compute:

- **the exact binomial probability,**

- **an approximation using the normal distribution for binomial (with continuity correction), and**

- **an approximation using the Poisson distribution.**

1.2 Part (a): Computations

Calculate and show the values of the probabilities in the 3-by-3 table below.

	Binomial Probability (exact)	Normal Approximation (with continuity correction)	Poisson Approximation
$k = 3$, $n = 6$, $p = 0.3$	0.1852	0.2015	0.1607
$k = 11$, $n = 40$, $p = 0.4$	0.0357	0.0353	0.0496
$k = 2$, $n = 400$, $p = 0.003$	0.2172	0.2746	0.2169

1.3 Part (b): Formulas

Provide the three formulas of the calculations you used in Part (a).

We calculate the exact probabilities using the PMF of the binomial distribution:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

We approximate the binomial probabilities using the Central Limit Theorem and the normal distribution, with continuity correction, according to the following formula:

$$P(X = k) \approx \Phi\left(\frac{k + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k - 0.5 - np}{\sqrt{np(1-p)}}\right).$$

Finally—in recognition of the fact that if $X_n \sim \mathrm{Bin}(n, p_n)$ with $np_n \to \lambda$ as $n \to \infty$, then $X_n \to_{\mathcal{D}} X \sim \mathrm{Pois}(\lambda)$ as $n \to \infty$—we approximate the binomial probabilities using the PMF of the Poisson distribution with rate parameter $\lambda = np$:

$$P(X = k) \approx \frac{(np)^k e^{-np}}{k!}.$$

1.4 Part (c): On Approximation Accuracy

Consider the following scenario: You are asked to approximate the binomial probability (which is too hard to calculate for some reason) in each of the two scenarios below by either normal approximations (with continuity correction) or Poisson approximation.

1. **Scenario (i): Approximate $P(X = 10)$ when $X \sim \mathrm{Bin}(30, \frac{1}{3})$.**
2. **Scenario (ii): Approximate $P(X = 3)$ when $X \sim \mathrm{Bin}(500, \frac{1}{1000})$.**

Which approximation method would you use? Your reason (in terms of sizes and relations of n, p, and k)? Your answer for this question need only be based on the approximation results in Part (a); no calculations are needed.

1. Under Scenario (i), I would prefer to use a normal approximation (with continuity correction) for $P(X = 10)$. Note that this scenario closely resembles the second row ("Case 2") of the table in Part (a): both have a moderate number of trials ($n = 40$ above; $n = 30$ here), a modest probability of success ($p = 0.4$ above; $p = \frac{1}{3}$ here), and a modest number whose density we are approximating ($k = 11$ above; $k = 10$ here). In Case 2 of the table, the normal approximation was significantly closer to the true binomial probability than the Poisson approximation; we would therefore expect the same in this Scenario (i).

2. Under Scenario (ii), I would prefer to use a Poisson approximation for $P(X = 10)$. Note that this scenario closely resembles the third row ("Case 3") of the table in Part (a): both have a large number of trials ($n = 400$ above; $n = 500$ here), a small probability of success ($p = 0.003$ above; $p = \frac{1}{1000}$ here), and a small number whose density we are approximating ($k = 2$ above; $k = 3$ here). In Case 3 of the table, the Poisson approximation was significantly closer to the true binomial probability than the normal approximation; we would therefore expect the same in this Scenario (ii).

In general, since the binomial distribution converges in distribution to the Poisson distribution as $n \to \infty$ (so long as $n \times p_n$ converges to a number), we would expect the Poisson approximation to do very well for large numbers of trials, like the $n = 500$ trials in Scenario (ii). In Scenario (i), with only $n = 30$ trials, n is not yet large enough for the Poisson approximation to be very good. However, $n = 30$ is (conventionally) just large enough for the Central Limit Theorem to be useful, making the normal approximation perform relatively better in Scenario (i).

2 Exercise 4: Coverage Probabilities of Binomial Asymptotic Confidence Intervals

2.1 Introduction

In this problem, we explore coverage probabilities of binomial asymptotic confidence intervals.

2.2 Part (a): Small Sample Size

Generate $k = 100$ draws of iid random variables from $\text{Bin}(n, p)$ for $n = 30$, $p = 0.1$. For each draw, obtain the MLE $\hat{p} = \frac{1}{n} \sum_{i=1}^{n} X_i$ and use the test level $\alpha = 0.05$ to compute:

- The Wald confidence interval for p (with endpoints $\hat{p} \pm z_{\alpha/2} \sqrt{\frac{1}{n}\hat{p}(1 - \hat{p})}$),

- **The Wilson confidence interval for p (with endpoints $\frac{z_{\alpha/2}^2+2n\hat{p}\pm z_{\alpha/2}\sqrt{z_{\alpha/2}^2+4n\hat{p}(1-\hat{p})}}{2(n+z_{\alpha/2}^2)}$, and**

- **The arcsine (converted) confidence interval for p you derived in Exercise 3 (using the variation stabilization method).**

In your simulation,

i. **What is the theoretical coverage probability, that is, the theoretical expected proportion of the k confidence intervals that should contain the true parameter $p = 0.1$?**

ii. **For each type of confidence interval, find the proportion of the k confidence intervals that actually contains the true parameter $p = 0.1$, based on your simulation.**

iii. **Compare and comment on the goodness of coverage probability of each of the three confidence intervals, based on your simulation results.**

First, recall that in Exercise 3, we found that the variance-stabilizing transformation (VST)-based confidence intervals for the probability of success parameter p of the binomial distribution have the form:

$$\left(\sin^2\left(\max\left\{\arcsin\sqrt{\hat{p}} - \frac{z_{1-\alpha/2}}{2\sqrt{n}}, 0\right\}\right), \ \sin^2\left(\min\left\{\arcsin\sqrt{\hat{p}} + \frac{z_{1-\alpha/2}}{2\sqrt{n}}, \frac{\pi}{2}\right\}\right)\right)$$

In the table below, we present the theoretical and empirical (in our specific simulation) coverage probabilities for each type of confidence interval.

	Wald	Wilson	VST
Theoretical Coverage Probability	0.95	0.95	0.95
Empirical Coverage Probability	0.88	0.99	0.97

In short, we would theoretically expect all three confidence interval types to have coverage probabilities of 0.95. In practice, while the Wilson and VST-based confidence intervals have coverage proportions close to (and even exceeding!) 0.95, the Wald coverage proportion of 0.88 falls noticeably short of 0.95. While the Wald, Wilson, and VST confidence intervals all appeal to the Central Limit Theorem for $\hat{p} = \bar{X}$, and therefore all build in a single approximation (since $n = 30$ is not ∞), the Wald confidence intervals use an additional approximation for the variance of the distribution of $\hat{p}$. The two stages of approximation involved with constructing Wald intervals explain why this coverage proportion falls short of 0.95, while the single stage of approximation involved with constructing Wilson and VST intervals allows their coverage proportions to be closer to the theoretical 0.95.

2.3 Part (b): Large Sample Size

Repeat the simulation in Part (a) with increased sample size $n = 150$. Do any of your previous conclusions for $n = 30$ change? Why?

In the table below, we present the new theoretical and empirical (in our specific simulation) coverage probabilities for each type of confidence interval, using our larger sample size of $n = 150$.

	Wald	Wilson	VST
Theoretical Coverage Probability	0.95	0.95	0.95
Empirical Coverage Probability	0.96	0.98	0.97

Now, all three types of confidence intervals have coverage proportions meeting (and even exceeding!) the theoretical 0.95 probability. By substantially increasing the number of trials in each simulation (from $n = 30$ to $n = 150$), the Wald confidence intervals are now able to overcome the error attributable to their "second stage" of approximation. It is also worth noting that the Wilson coverage proportion has decreased from 0.99 to 0.98—with a larger number of trials built into each simulation, it is more difficult to achieve the exceptionally high coverage proportion that occurred by chance before. The VST coverage proportions have stayed the same, at 0.97, by chance.