24500 HW1

Bin Yu

January 17, 2025

Question 1

Let X be a Poisson random variable with parameter λ,

$$P(X = k) \; = \; \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \dots$$

$$E[X] = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!}.$$

Factor out λ from the summation by $k \cdot \frac{\lambda^k}{k!} = \lambda \cdot \frac{\lambda^{k-1}}{(k-1)!}$. Thus,

$$E[X] = \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1} e^{-\lambda}}{(k-1)!} = \lambda \, e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda \, e^{-\lambda} \, e^{\lambda} = \lambda.$$

$$E[X] = \lambda.$$

For $E[X^2]$:

$$E[X^2] = \sum_{k=0}^{\infty} k^2 \cdot \frac{\lambda^k e^{-\lambda}}{k!}.$$

We can rewrite $k^2 = k(k-1) + k$, and then split the sum:

$$E[X^2] = \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k e^{-\lambda}}{k!} \; + \; \sum_{k=0}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!}.$$

We have

$$\sum_{k=0}^{\infty} k(k-1) \, \frac{\lambda^k}{k!} = \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2 \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda^2 e^{\lambda}$$

and similarly

$$\sum_{k=0}^{\infty} k \, \frac{\lambda^k}{k!} = \lambda e^{\lambda}.$$

Multiplying by $e^{-\lambda}$ in each case:

$$E[X^2] = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda e^{-\lambda} e^{\lambda} = \lambda^2 + \lambda.$$

$$E[X^2] = \lambda^2 + \lambda.$$

For $E(X^3)$:

$$X^3 = X(X-1)(X-2) + 3X^2 - 2X.$$

Taking expectations on both sides:

$$E[X^3] = E\big[X(X-1)(X-2)\big] + 3\,E[X^2] - 2\,E[X].$$

We have:

$$E[X(X-1)(X-2)] \;=\; \sum_{k=0}^{\infty} k(k-1)(k-2)\; P(X=k).$$

Observe that $k(k-1)(k-2) = 0$ for $k = 0, 1, 2$. The sum starts at $k = 3$:

$$E[X(X-1)(X-2)] = \sum_{k=3}^{\infty} k(k-1)(k-2)\,\frac{\lambda^k e^{-\lambda}}{k!}.$$

$$E[X(X-1)(X-2)] = e^{-\lambda} \sum_{k=3}^{\infty} \frac{\lambda^k}{(k-3)!}. = e^{-\lambda}\,\lambda^3 \sum_{k=3}^{\infty} \frac{\lambda^{k-3}}{(k-3)!}.$$

Let $j = k - 3$. Then j goes from 0 to ∞,

$$E[X(X-1)(X-2)] = e^{-\lambda}\,\lambda^3 \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = e^{-\lambda}\,\lambda^3\,e^{\lambda} = \lambda^3.$$

$$E\big[X(X-1)(X-2)\big] = \lambda^3.$$

$$E[X^3] = E\big[X(X-1)(X-2)\big] + 3\,E[X^2] - 2\,E[X].$$

Substitute the values of each term:

$$E[X^3] = \lambda^3 + 3(\lambda^2 + \lambda) - 2\lambda = \lambda^3 + 3\lambda^2 + 3\lambda - 2\lambda = \lambda^3 + 3\lambda^2 + \lambda.$$

Therefore,

$$E[X^3] = \lambda^3 + 3\lambda^2 + \lambda.$$

For $E(X^4)$:

$$X^4 = X(X-1)(X-2)(X-3) \;+\; 6X^3 \;-\; 11X^2 \;+\; 6X.$$

Taking expectations:

$$E[X^4] = E\big[X(X-1)(X-2)(X-3)\big] + 6\,E[X^3] - 11\,E[X^2] + 6\,E[X].$$

$$E[X(X-1)(X-2)(X-3)] \;=\; \sum_{k=0}^{\infty} k(k-1)(k-2)(k-3)\; P(X=k).$$

Since $k(k-1)(k-2)(k-3) = 0$ for $k = 0, 1, 2, 3$, the sum starts at $k = 4$, therefore,

$$E[X(X-1)(X-2)(X-3)] = e^{-\lambda} \sum_{k=4}^{\infty} \frac{\lambda^k}{(k-4)!} = e^{-\lambda}\,\lambda^4 \sum_{k=4}^{\infty} \frac{\lambda^{k-4}}{(k-4)!}.$$

Let $j = k - 4$. Then j goes from 0 to ∞, and

$$E[X(X-1)(X-2)(X-3)] = e^{-\lambda}\,\lambda^4 \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = e^{-\lambda}\,\lambda^4\,e^{\lambda} = \lambda^4.$$

Thus,

$$E\big[\,X(X-1)(X-2)(X-3)\big] = \lambda^4.$$

We already have

$$E[X^3] = \lambda^3 + 3\lambda^2 + \lambda, \quad E[X^2] = \lambda^2 + \lambda, \quad E[X] = \lambda.$$

Substitute each term:

$$\begin{aligned} E[X^4] &= \lambda^4 + 6(\lambda^3 + 3\lambda^2 + \lambda) - 11(\lambda^2 + \lambda) + 6\lambda \\ &= \lambda^4 + 6\lambda^3 + 18\lambda^2 + 6\lambda - 11\lambda^2 - 11\lambda + 6\lambda \\ &= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda. \end{aligned}$$

Therefore,

$$\boxed{E[X^4] = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda.}$$

In conclusion,

$$E[X] = \lambda, \quad E[X^2] = \lambda^2 + \lambda, \quad E[X^3] = \lambda^3 + 3\lambda^2 + \lambda, \quad E[X^4] = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda.$$

Question 2

$$f(x) \;=\; \begin{cases} \lambda\,e^{-\lambda x}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

$$E[X] \;=\; \int_0^{\infty} x\,f(x)\,dx \;=\; \int_0^{\infty} x\,\lambda\,e^{-\lambda x}\,dx.$$

Let

$$u = x \quad \Longrightarrow \quad du = dx, \qquad dv = \lambda e^{-\lambda x}\,dx \quad \Longrightarrow \quad v = -\,e^{-\lambda x}.$$

$$\int_0^{\infty} x\,\lambda\,e^{-\lambda x}\,dx = \big(x \cdot -e^{-\lambda x}\big)\Big|_0^{\infty} \;-\; \int_0^{\infty} \big(-e^{-\lambda x}\big)\,dx.$$

As $x \to \infty$, $x\,e^{-\lambda x} \to 0$. At $x = 0$, $x\,e^{-\lambda x} = 0$.

$$-xe^{-\lambda x}\Big|_0^{\infty} \;=\; 0.$$

Therefore,

$$0 - \int_0^{\infty} \big(-e^{-\lambda x}\big)\,dx \;=\; \int_0^{\infty} e^{-\lambda x}\,dx.$$

$$\int_0^{\infty} e^{-\lambda x}\,dx = \Big[-\frac{1}{\lambda}\,e^{-\lambda x}\Big]_0^{\infty} = 0 \;-\; \big(-\tfrac{1}{\lambda}\big) = \tfrac{1}{\lambda}.$$

$$\boxed{E[X] \;=\; \frac{1}{\lambda}.}$$

Question 3

A random variable $X \sim \text{Binomial}(n, p)$ satisfies:

$$P(X = k) \;=\; \binom{n}{k}\, p^k \,(1-p)^{\,n-k}, \qquad k = 0, 1, \ldots, n.$$

This is the *exact* probability.

When n is large and p is neither too close to 0 nor 1, based on CLT, we can approximate X by a Normal random variable $Y \sim \mathcal{N}(\mu, \sigma^2)$ where

$$\mu = np, \quad \sigma^2 = np(1-p).$$

To approximate $P(X = k)$, use the continuity correction, i.e.

$$P(X = k) \;\approx\; \Phi\Big(\frac{k + 0.5 - \mu}{\sigma}\Big) \;-\; \Phi\Big(\frac{k - 0.5 - \mu}{\sigma}\Big),$$

where Φ is the CDF of the standard Normal distribution.

When n is large and p is small, that np converge to λ, the Binomial random variable X can be approximated by a Poisson random variable with parameter $\lambda = np$. That is:

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

(a) $n = 7$, $p = 0.3$, $k = 3$

Exact Binomial:

$$P_{\text{Bin}}(X = 3) = \binom{7}{3}\,(0.3)^3\,(0.7)^4 = 35 \times 0.027 \times 0.2401 \;\approx\; 0.226894.$$

Normal Approx. (with continuity):

$$\mu = 2.1, \quad \sigma \approx 1.21.$$

$$z_1 = \frac{3.5 - 2.1}{1.21} \approx 1.157, \quad z_2 = \frac{2.5 - 2.1}{1.21} \approx 0.331.$$

$$P_{\text{Normal}}(X = 3) \approx \Phi(z_1) - \Phi(z_2) = 0.249804.$$

Poisson Approx.:

$$\lambda = np = 2.1, \quad P_{\text{Pois}}(X = 3) = e^{-2.1}\,\frac{(2.1)^3}{3!}.$$

$$(2.1)^3 = 9.261, \quad 3! = 6, \quad \frac{(2.1)^3}{6} \approx 1.5435, \quad e^{-2.1} \approx 0.12246,$$

thus

$$P_{\text{Pois}}(X = 3) \approx 1.5435 \times 0.12246 \approx 0.189011.$$

Comment: Since n is quite small, it doesn't fulfill the condition of using CLT, the Normal approximation is not accurate. Poisson is not accurate (since p is not small enough, which means that it is not a rare case).

(b) $n = 40,\ p = 0.4,\ k = 11$

Exact Binomial:

$$P_{\text{Bin}}(X = 11) = \binom{40}{11} (0.4)^{11} (0.6)^{29}.$$

$$P_{\text{Bin}}(X = 11) \approx 0.035727.$$

Normal Approx. (with continuity):

$$\mu = 16, \quad \sigma = \sqrt{9.6} \approx 3.098.$$

$$z_1 = \frac{11.5 - 16}{3.098} \approx -1.45, \quad z_2 = \frac{10.5 - 16}{3.098} \approx -1.77.$$

$$P_{\text{Normal}}(X = 11) \approx \Phi(z_1) - \Phi(z_2) = 0.035018.$$

Poisson Approx.:

$$\lambda = 16, \quad P_{\text{Pois}}(X = 11) = e^{-16} \frac{16^{11}}{11!}.$$

$$P_{\text{Pois}}(X = 11) \approx 0.049597.$$

Comment: Here n is bigger than 20 and $p = 0.4$ is not too small, so the Normal approximation is quite accurate, its value (≈ 0.035018) is close to the exact Binomial (≈ 0.035727). This is because $np = 16$, it can fulfill the assumption of CLT that $np \to \infty$. The Poisson approximation is not accurate (≈ 0.049597) because p is large here, and $\lambda = 16$ is relatively large, which means that it is not a rare case, and should not be modeled as poission distribution.

(c) $n = 400,\ p = 0.0025,\ k = 2$

Exact Binomial:

$$P_{\text{Bin}}(X = 2) = \binom{400}{2} (0.0025)^2 (0.9975)^{398}.$$

$$P_{\text{Bin}}(X = 2) \ \approx \ 0.184170.$$

Normal Approx. (with continuity):

$$\mu = 1, \quad \sigma^2 = 0.9975, \quad \sigma \approx 0.9987.$$

$$z_1 = \frac{2.5 - 1}{0.9987} \approx 1.50, \quad z_2 = \frac{1.5 - 1}{0.9987} \approx 0.50.$$

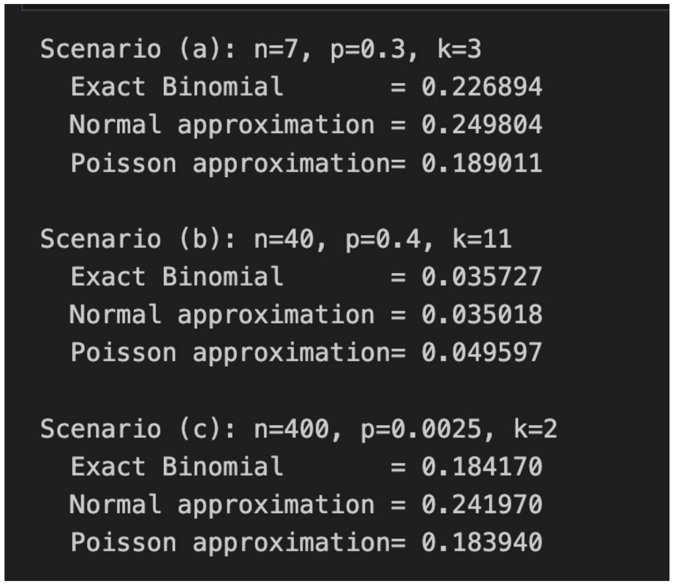
$$P_{\text{Normal}}(X = 2) \approx \Phi(1.50) - \Phi(0.50) = 0.241970.$$

Poisson Approx.:

$$\lambda = 1, \quad P_{\text{Pois}}(X = 2) = e^{-1} \frac{1^2}{2!} = \frac{e^{-1}}{2} \approx 0.183940.$$

Comment: Because n is large and p is very small, means that it is a rare case and $np \to \lambda$. So the Poisson approximation (≈ 0.183940) is accurate and almost matches the exact Binomial (≈ 0.184170). The Normal approximation is not accurate (around 0.241970), since the assumption of CLT that $np \to \infty$ is not satisfied, where $np = 1$.

Python Approximation



```
Scenario (a): n=7, p=0.3, k=3
  Exact Binomial        = 0.226894
  Normal approximation = 0.249804
  Poisson approximation= 0.189011

Scenario (b): n=40, p=0.4, k=11
  Exact Binomial        = 0.035727
  Normal approximation = 0.035018
  Poisson approximation= 0.049597

Scenario (c): n=400, p=0.0025, k=2
  Exact Binomial        = 0.184170
  Normal approximation = 0.241970
  Poisson approximation= 0.183940
```

Figure 1: Python Approximation

Question 4

(a)

Let $X_1, X_2, \ldots, X_n$ be i.i.d. Bernoulli(p). The likelihood function is:

$$L(p) = \prod_{i=1}^{n} p^{X_i} (1-p)^{1-X_i}.$$

the log-likelihood is

$$\ell(p) = \sum_{i=1}^{n} \big[X_i \ln p + (1-X_i) \ln(1-p)\big].$$

$$\frac{d}{dp}\,\ell(p) = \frac{d}{dp}\left(\sum_{i=1}^{n} \big[X_i \ln(p) + (1-X_i)\ln(1-p)\big]\right).$$

$$\frac{d}{dp}\,\ell(p) = \sum_{i=1}^{n} \left[\frac{X_i}{p} - \frac{1-X_i}{1-p}\right].$$

$$\frac{d}{dp}\,\ell(p) = \sum_{i=1}^{n} \frac{X_i(1-p) - p(1-X_i)}{p(1-p)}.$$

$$\frac{d}{dp}\,\ell(p) = \sum_{i=1}^{n} \frac{X_i - p}{p(1-p)}.$$

Set the derivative to zero and solve for p.

$$\frac{1}{p(1-p)}\ \sum_{i=1}^{n}(X_i - p)\ =\ 0.$$

Since $p(1-p) \neq 0$ for $p \in (0,1)$,

$$\sum_{i=1}^{n}\big(X_i - p\big) = 0$$

$$\sum_{i=1}^{n} X_i - np = 0$$

$$p = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Therefore,

$$\boxed{\widehat{p}_{\text{MLE}} = \frac{1}{n}\sum_{i=1}^{n} X_i.}$$

(b)

Since each X_i is Bernoulli(p),

$$E[\,X_i\,] = p, \quad \text{Var}(X_i) = p(1-p).$$

$$E[\widehat{p}_{\text{MLE}}] = E\Big[\frac{1}{n}\sum X_i\Big] = \frac{1}{n}\cdot n \cdot E[\,X_i\,] = p,$$

$$\text{Var}(\widehat{p}_{\text{MLE}}) = \text{Var}\Big(\frac{1}{n}\sum X_i\Big) = \frac{1}{n^2}\,n\,p(1-p) = \frac{p(1-p)}{n}.$$

For MSE:

$$\text{MSE}(\hat{\theta}) = E\big[(\hat{\theta}-\theta)^2\big] = E\Big[\big((\hat{\theta}-E[\hat{\theta}])\ +\ (E[\hat{\theta}]-\theta)\big)^2\Big]$$

$$(\hat{\theta}-\theta)^2\ =\ (\hat{\theta}-E[\hat{\theta}])^2\ +\ (E[\hat{\theta}]-\theta)^2\ +\ 2\,(\hat{\theta}-E[\hat{\theta}])\,\big(E[\hat{\theta}]-\theta\big).$$

$$\text{MSE}(\hat{\theta})\ =\ E\big[(\hat{\theta}-E[\hat{\theta}])^2\big]\ +\ E\big[(E[\hat{\theta}]-\theta)^2\big]\ +\ 2\,E\Big[(\hat{\theta}-E[\hat{\theta}])\,(E[\hat{\theta}]-\theta)\Big].$$

Since:

$$E\big[(E[\hat{\theta}]-\theta)^2\big] = (E[\hat{\theta}]-\theta)^2.$$

and,

$$E\Big[(\hat{\theta}-E[\hat{\theta}])\,(E[\hat{\theta}]-\theta)\Big]\ =\ (E[\hat{\theta}]-\theta)\,E\big[\hat{\theta}-E[\hat{\theta}]\big]\ =\ (E[\hat{\theta}]-\theta)\,(E[\hat{\theta}]-E[\hat{\theta}])\ =\ 0.$$

$$\text{MSE}(\hat{\theta})\ =\ E\big[(\hat{\theta}-E[\hat{\theta}])^2\big]\ +\ (E[\hat{\theta}]-\theta)^2.$$

Since $E\big[(\hat{\theta}-E[\hat{\theta}])^2\big] = \text{Var}(\hat{\theta})$, and $E[\hat{\theta}]-\theta$ is Bias($\hat{\theta}$):

$$\mathrm{MSE}(\hat{\theta}) \;=\; \mathrm{Var}(\hat{\theta}) \;+\; \bigl(\mathrm{Bias}(\hat{\theta})\bigr)^2.$$

Therefore,

$$\mathrm{MSE}\bigl(\widehat{p}_{\mathrm{MLE}}\bigr) = E\bigl[(\widehat{p}_{\mathrm{MLE}} - p)^2\bigr] = \mathrm{Var}(\widehat{p}_{\mathrm{MLE}}) + \bigl(\mathrm{Bias}\bigr)^2.$$

Since $\widehat{p}_{\mathrm{MLE}}$ is unbiased, bias $= 0$,

$$\mathrm{MSE}(\widehat{p}_{\mathrm{MLE}}) = \frac{p(1-p)}{n}.$$

(c)

By the Central Limit Theorem:

$$\widehat{p}_{\mathrm{MLE}} \;\approx\; N\Bigl(p,\; \frac{p(1-p)}{n}\Bigr) \quad \text{for large } n.$$

or

$$\sqrt{n}\,\bigl(\,\widehat{p}_{\mathrm{MLE}} - p\,\bigr) \approx N\bigl(0,\, p(1-p)\bigr).$$

(d)

From the asymptotic theory (or the Central Limit Theorem), we know:

$$\widehat{p}_{\mathrm{MLE}} \;\approx\; \mathcal{N}\Bigl(p,\; \frac{p(1-p)}{n}\Bigr) \quad \text{for large } n.$$

By the Weak Law of Large Numbers (WLLN), $\widehat{p} \to p$ in probability. Moreover, the Central Limit Theorem shows:

$$\sqrt{n}\,\frac{\widehat{p} - p}{\sqrt{p\,(1-p)}} \quad \to \quad N(0,1) \quad (\text{as } n \to \infty).$$

However, we do not know p in practice. Under Slutsky's theorem, since $\widehat{p} \to p$, we can replace p by $\widehat{p}$ in the denominator,

$$\sqrt{n}\,\frac{\widehat{p} - p}{\sqrt{\widehat{p}\,(1-\widehat{p})}} \quad \to \quad N(0,1) \quad (\text{as } n \to \infty).$$

$$P\Bigl(-\,z_{1-\alpha/2} \;\le\; \sqrt{n}\,\frac{\widehat{p} - p}{\sqrt{\widehat{p}\,(1-\widehat{p})}} \;\le\; z_{1-\alpha/2}\Bigr) \;\approx\; 1-\alpha.$$

$$-\,z_{1-\alpha/2} \;\le\; \sqrt{n}\,\frac{\widehat{p} - p}{\sqrt{\widehat{p}\,(1-\widehat{p})}} \;\le\; z_{1-\alpha/2}$$

$$\widehat{p} - z_{1-\alpha/2}\sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}} \;\le\; p \;\le\; \widehat{p} + z_{1-\alpha/2}\sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{n}}$$

Therefore, the Wald $(1-\alpha)$-level confidence interval for p is:

$$\left[\widehat{p} - z_{1-\alpha/2}\sqrt{\tfrac{\widehat{p}(1-\widehat{p})}{n}}\ ,\ \widehat{p} + z_{1-\alpha/2}\sqrt{\tfrac{\widehat{p}(1-\widehat{p})}{n}}\ \right].$$

For 95% CI, $z_{0.975} \approx 1.96$. The 95% CI is given by:

$$\left[\widehat{p} - 1.96\sqrt{\tfrac{\widehat{p}(1-\widehat{p})}{n}}\ ,\ \widehat{p} + 1.96\sqrt{\tfrac{\widehat{p}(1-\widehat{p})}{n}}\ \right].$$

(e)

Starting from:

$$\widehat{p} - z_{1-\alpha/2}\sqrt{\frac{p\,(1-p)}{n}} \ \le\ p \ \le\ \widehat{p} + z_{1-\alpha/2}\sqrt{\frac{p\,(1-p)}{n}}$$

$$\left|\hat{p} - p\right| \ \le\ z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}},$$

denote $z = z_{1-\alpha/2}$. Squaring both sides:

$$(\hat{p} - p)^2 \ =\ z^2\,\frac{p\,(1-p)}{n}.$$

$$\hat{p}^2 \ -\ 2\hat{p}\,p \ +\ p^2 \ =\ \frac{z^2}{n}\,p \ -\ \frac{z^2}{n}\,p^2$$

$$p^2 \ +\ \frac{z^2}{n}\,p^2 \ -\ 2\,\hat{p}\,p \ -\ \frac{z^2}{n}\,p \ +\ \hat{p}^2 \ =\ 0.$$

$$\left(1 + \tfrac{z^2}{n}\right)p^2 \ -\ \left(2\,\hat{p} + \tfrac{z^2}{n}\right)p \ +\ \hat{p}^2 \ =\ 0.$$

let

$$a = 1 + \frac{z^2}{n}, \quad b = -\left(2\,\hat{p} + \frac{z^2}{n}\right), \quad c = \hat{p}^2.$$

The general solution of $a\,p^2 + b\,p + c = 0$ is

$$p = \frac{-b \pm \sqrt{b^2 - 4\,a\,c}}{2\,a}.$$

Substitute a,b,c:

$$p = \frac{\left(2\,\hat{p} + \tfrac{z^2}{n}\right) \pm \sqrt{\left(2\,\hat{p} + \tfrac{z^2}{n}\right)^2 \ -\ 4\left(1 + \tfrac{z^2}{n}\right)\hat{p}^2}}{2\left(1 + \tfrac{z^2}{n}\right)}.$$

$$p = \frac{\left(2\,\hat{p} + \tfrac{2z^2}{2n}\right) \pm \sqrt{\left(2\,\hat{p} + \tfrac{2z^2}{2n}\right)^2 \ -\ 4\left(1 + \tfrac{z^2}{n}\right)\hat{p}^2}}{2\left(1 + \tfrac{z^2}{n}\right)}.$$

$$p = \frac{2\left(\hat{p} + \tfrac{z^2}{2n}\right) \pm 2\sqrt{\left(\hat{p} + \tfrac{z^2}{2n}\right)^2 \ -\ \left(1 + \tfrac{z^2}{n}\right)\hat{p}^2}}{2\left(1 + \tfrac{z^2}{n}\right)}.$$

$$p = \frac{\left(\hat{p} + \frac{z^2}{2n}\right) \pm \sqrt{\hat{p}^2 + \frac{z^2}{n}\hat{p} + (\frac{z^2}{2n})^2 \; - \; \hat{p}^2 - \frac{z^2}{n}\hat{p}^2}}{\left(1 + \frac{z^2}{n}\right)}.$$

$$p = \frac{\hat{p} + \frac{z^2}{2n} \pm \sqrt{\frac{z^2}{n}(\hat{p} - \hat{p}^2) + (\frac{z^2}{2n})^2}}{\left(1 + \frac{z^2}{n}\right)}.$$

$$\boxed{p = \frac{\hat{p} + \frac{z^2}{2n} \; \pm \; z\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}}.}$$

For 95% CI, the $z_{0.975} \approx 1.96$, the CI is given by:

$$\left[\frac{\hat{p} \; + \; \dfrac{(1.96)^2}{2n} \; - \; 1.96\sqrt{\dfrac{\hat{p}\,(1-\hat{p})}{n} \; + \; \dfrac{(1.96)^2}{4n^2}}}{1 + \dfrac{(1.96)^2}{n}} \;,\; \frac{\hat{p} \; + \; \dfrac{(1.96)^2}{2n} \; + \; 1.96\sqrt{\dfrac{\hat{p}\,(1-\hat{p})}{n} \; + \; \dfrac{(1.96)^2}{4n^2}}}{1 + \dfrac{(1.96)^2}{n}}\right]$$

(f)

Visualize the Wilson's method, to plot

$$(\hat{p}_{\text{MLE}} - p)^2 \quad \text{and} \quad R^2\,p\,(1-p)$$

Where,

- $\hat{p}_{\text{MLE}}$ is a chosen sample proportion (e.g. 0.5 or 0.3),
- R^2 is a constant equals to $\frac{z^2}{n}$

Below are three examples of such plots:

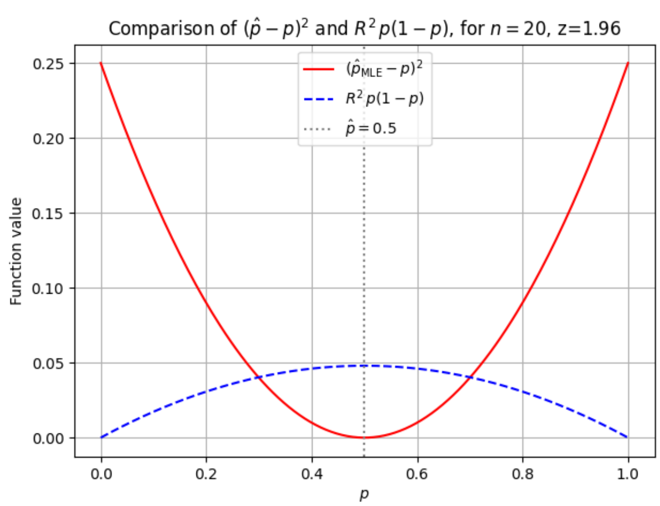


Figure 2: Comparison of $(\hat{p}_{\text{MLE}} - p)^2$ (red) and $R^2\,p(1-p)$ (blue), here with $n = 20$, $z = 1.96$, and (for example) $\hat{p}_{\text{MLE}} = 0.5$.

In Figure 2, the red curve is $(\hat{p}_{\text{MLE}} - p)^2$ with $\hat{p}_{\text{MLE}} = 0.5$, which has its minimum at $p = 0.5$. The blue dashed curve is $R^2\, p(1-p)$ with $R^2 = \frac{z^2}{n}$. The intersection points between these curves are the p_{left} and p_{right} that solve

$$(\hat{p}_{\text{MLE}} - p)^2 = z^2\, \frac{p(1-p)}{n}.$$

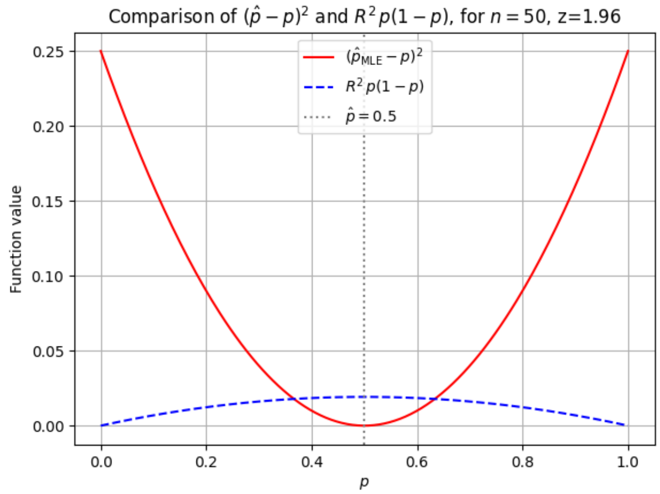


Figure 3: Comparison for $n = 50$, $z = 1.96$, $\hat{p}_{\text{MLE}} = 0.5$.

Figure 3 repeats the idea but for $n = 50$, $z = 1.96$. Notice the peak of the blue curve ($R^2\, p(1-p)$) is lower because the factor $R^2 = \frac{(1.96)^2}{50}$ is smaller than $\frac{(1.96)^2}{20}$; Therefore, as n increases, the resulting interval becomes smaller.

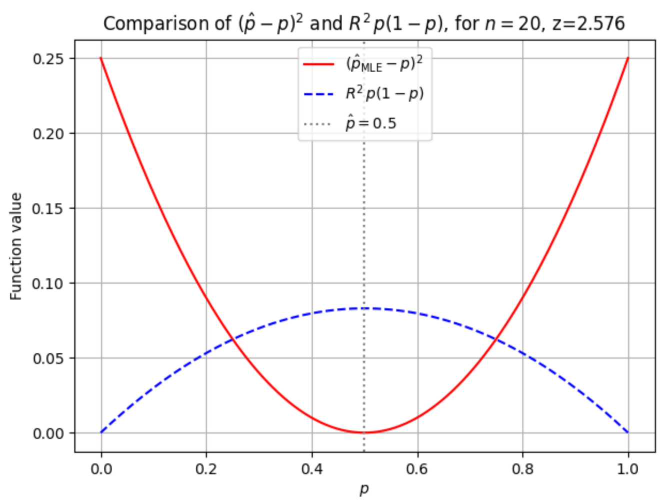


Figure 4: Comparison for $n = 20$, but $z = 2.576$ for 99% CI.

Finally, Figure 4 shows the same $n = 20$ but with $z = 2.576$ (approximately for a 99% confidence level). The dashed curve is taller (since $R^2 = \frac{(2.576)^2}{20}$ is larger than $\frac{(1.96)^2}{20}$). As a result, as $1 - a$ increases, the intersection points are farther away from $\hat{p}_{\text{MLE}} = 0.5$, giving a wider interval.

Interpretation of p_{left} and p_{right}: They are the left and right bounds of the Wilson interval. In each figure, a larger n or smaller z (correspond to smaller $1 - a$) reduces the peak of the blue curve and yields an interval closer to $\hat{p}_{\text{MLE}}$ (smaller), whereas a larger z (correspond to larger $1-a$) or smaller n raises that curve and yields a wider range of p.

Another property of left and right bounds is that as we can see from the graph above, they are never negative.

(g)

Case 1: $\hat{p}_{\text{MLE}} = 0.5$

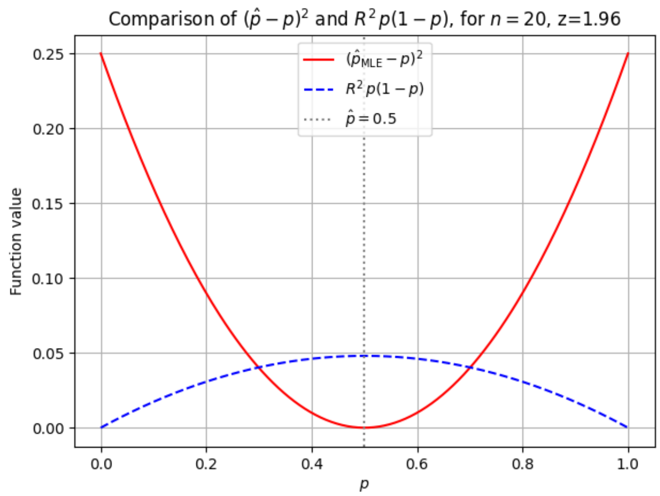


Figure 5: $\widehat{p}_{MLE} = 0.5$

In Figure 5, the minimum of $(\hat{p}_{\text{MLE}} - p)^2$ is at $p = 0.5$, exactly matching $\hat{p}_{\text{MLE}}$. The two curves meet at points symmetrically around 0.5, giving a symmetric interval that $\hat{p}_{\text{MLE}} - p_{\text{left}} = p_{\text{right}} - \hat{p}_{\text{MLE}}$.

Case 2: $\hat{p}_{\text{MLE}} = 0.75$

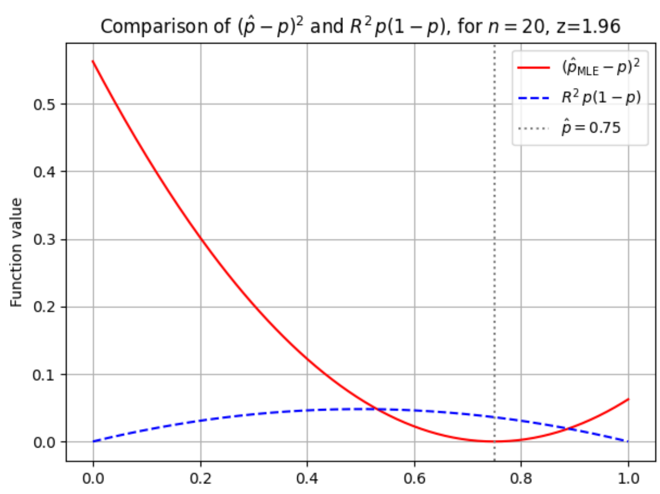


Figure 6: $\hat{p}_{\text{MLE}} = 0.75$

Here, in Figure 6, the red curve $(\hat{p}_{\text{MLE}} - p)^2$ attains its minimum at $p = 0.75$. The left intersection point with the blue curve will be further away from 0.75 than the right intersection, reflecting an asymmetry. (i.e. $\hat{p}_{\text{MLE}} - p_{\text{left}}$ is bigger than $p_{\text{right}} - \hat{p}_{\text{MLE}}$)

Case 3: $\hat{p}_{\text{MLE}} = 0.25$

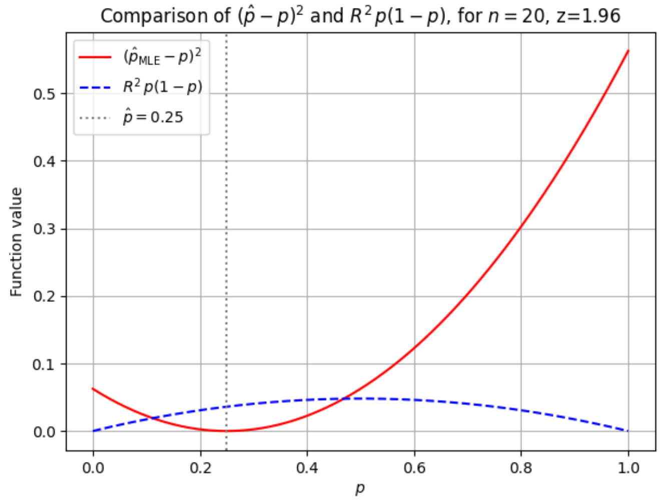


Figure 7: $\hat{p}_{\text{MLE}} = 0.25$

In Figure 7, we see the red curve's minimum occurs at $p = 0.25$. The intersection on the right side is farther from 0.25 than the left intersection (i.e. $\hat{p}_{\text{MLE}} - p_{\text{left}}$ is smaller than $p_{\text{right}} - \hat{p}_{\text{MLE}}$).

Interpretation:

If $\hat{p}_{\text{MLE}}$ is not at 0.5, the interval is often not symmetric about $\hat{p}_{\text{MLE}}$. This is reasonable with the skewness of the binomial distribution in these cases, and its error distribution should be skewed as well.