STAT201A: Introduction to Probability at an Advanced Level (Fall 2024)
UC Berkeley

Problem Set 3

Due: 10:00pm, Friday, October 11, 2024 (via Gradescope)

1. **(Approximating Binomial Distributions)** The goal of this question is to empirically verify three approximations to the exact Binomial probability $\mathbb{P}(X = k)$, where $X \sim \text{Binomial}(n, p)$:

   - $\mathbb{P}(Y = k)$, where $Y \sim \text{Poisson}(np)$, the Poison approximation with rate parameter np;
   - The normal approximation
   $$\phi(k; np, np(1-p)) := \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left\{-\frac{(k-np)^2}{2np(1-p)}\right\}$$
   - The entropic approximation
   $$\text{Ent}(k; n, p) := \frac{1}{\sqrt{2\pi nf(1-f)}} \exp(-n\text{KL}(f\|p))$$
   where $f = \frac{k}{n}$ and $\text{KL}(f\|p) = f \log \frac{f}{p} + (1-f) \log \frac{1-f}{1-p}$.

   (a) Take $n = 30$ and $p = 0.05$. Create a table (31 rows and 3 columns) containing the absolute errors for each approximation,
   $$|\mathbb{P}(X = k) - \mathbb{P}(Y = k)|, \quad |\mathbb{P}(X = k) - \phi(k; np, np(1-p))|$$
   and
   $$|\mathbb{P}(X = k) - \text{Ent}(k; n, p)|$$
   for $k = 0, 1, \ldots, 30$. (Note: The entropic approximation does not exist for $k = 0$ and $k = 30$, so only list it for $k = 1, \ldots, 29$).
   Based on the table, comment on the accuracy of each of the three approximations for the Binomial distribution.

   (b) Create a similar table for the relative errors
   $$\frac{|\mathbb{P}(X = k) - \mathbb{P}(Y = k)|}{\mathbb{P}(X = k)}, \quad \frac{|\mathbb{P}(X = k) - \phi(k; np, np(1-p))|}{\mathbb{P}(X = k)}$$
   and
   $$\frac{|\mathbb{P}(X = k) - \text{Ent}(k; n, p)|}{\mathbb{P}(X = k)}$$
   for $k = 0, 1, \ldots, 30$. Based on this table, comment on the accuracy of each of the three approximations for the Binomial.

   (c) Repeat exercises (a) and (b) above for $n = 30$ and $p = 0.25$.

   (d) Repeat exercises (a) and (b) above for $n = 30$ and $p = 0.5$.

   See attached PDF.

2. **(KL-Divergence, Multinomial)** Let X and Y be discrete random variables with distributions p and q respectively (So $p(k) = \mathbb{P}(X = k)$ and $q(k) = \mathbb{P}(Y = k)$). Remember that the Kullback–Leibler divergence is defined by

$$\mathrm{KL}(p\|q) := \mathbb{E}_p\Big[\ln\big(\frac{p(X)}{q(X)}\big)\Big] = \sum_k p(k)\ln\big(\frac{p(k)}{q(k)}\big).$$

(a) Show that when $q(k)$ is a Poisson distribution with parameter $\lambda > 0$, then the KL-divergence is minimized by setting λ to be the mean of $p(k)$.

(b) Rememebr that the entropy $H(p)$ is defined to be $H(p|q) := -\mathbb{E}_p[\ln(p(X))]$. Assume that we need to place n balls into d bins. The number of ways to place the balls resulting in k_i total balls in bin i, for $i = 1, \dots, d$, is given by the combinatorial expression $\binom{n}{k_1,k_2,\dots,k_d}$. Now consider the empirical distribution of the balls. Its probability mass function is $p(i) = k_i/n$. Let N_p denote the number of configurations with empirical distribution p, show that

$$\ln(N_p) = nH(p) + O(\ln(n)),$$

where $h(p)$ is the entropy of p.

In other words, there are many more high-entropy configurations than low-entropy configurations. This suggests the intuition that, if we consider a physical system at a "macro level" (such as the distribution of gas particles in a container) then we should expect it to drift toward high-entropy configurations.

Hint: Recall Stirling's approximation

$$\ln(n!) = n\ln(n) - n + O(\ln(n)).$$

(a) if $q(k)$ is Poisson with parameter λ, then $q(k) = e^{-\lambda}\lambda^k/k!$. Let's calculate now the KL-divergence.

$$\begin{aligned}
\mathrm{KL}(p\|q) &= \sum_k p(k)\ln\big(\frac{p(k)}{q(k)}\big) \\
&= \sum_k p(k)\ln\big(p(k)\big) - \sum_k p(k)\ln\big(e^{-\lambda}\lambda^k/k!\big) \\
&= \sum_k p(k)\ln\big(p(k)\big) + \lambda\sum_k p(k) - \sum_k p(k)\Big(k\ln\big(\lambda\big) - \ln\big(k!\big)\Big) \\
&= \sum_k p(k)\ln\big(p(k)\big) + \lambda - \ln\big(\lambda\big)\sum_k kp(k) + \sum_k p(k)\ln\big(k!\big)
\end{aligned}$$

Taking the derivative we can verify that there is a minimum precisely when $\lambda = \sum_k kp(k)$.

(b) As stated at the beginning of the problem, $N_p = \frac{n!}{k_1!k_2!\cdots k_d!}$, we can now use Stirling approximation and do a calculation.

$$
\begin{aligned}
\log(N_p) &= \log(n!) - \sum_{i=1}^{d} \log(k_i!) \\
&= n\log(n) - n + O(\log n) - \Big(\sum_{i=1}^{d} k_i \log(k_i) - k_i + O(\log k_i)\Big) \\
&= n\log(n) - n + O(\log n) - \Big(\sum_{i=1}^{d} k_i \log(k_i) - k_i + O(\log n)\Big) \\
&= n\log(n) - n - \Big(\sum_{i=1}^{d} k_i \log(k_i) - k_i\Big) + O(\log n) \\
&= \Big(\sum_{i=1}^{d} k_i\Big)\log(n) - \Big(\sum_{i=1}^{d} k_i \log(k_i)\Big) + \Big(\sum_{i=1}^{d} k_i\Big) - n + O(\log n) \\
&= -\Big(\sum_{i=1}^{d} k_i\big(\log(k_i) - \log(n)\big)\Big) + O(\log n) \\
&= -n\Big(\sum_{i=1}^{d} \frac{k_i}{n}\log\big(\frac{k_i}{n}\big)\Big) + O(\log n) \\
&= nh(p) + O(\log n)
\end{aligned}
$$

The relevant identities we used are that $k_i \leq n$ and hence $\log(k_i) = O(\log n)$, that $\sum_{i=1}^{d} k_i = n$ and since d is fixed and finite, $O(\log n)\sum_{i=1}^{d} O(\log n)$ is still $O(\log n)$.

3. **(Poisson)** Let $K = X_1 + X_2 + \cdots + X_N$, where $N \sim \text{Poisson}(\lambda)$ and $X_1, X_2, \cdots$ are independent Bernoulli (p) random variables. Assuming that N and $\{X_i\}_{i\in\mathbb{N}}$ are mutually independent, find the distribution of K.

$$M_N(t) = \sum_{i=0}^{\infty} e^{-\lambda}\frac{\lambda^i}{i!}e^{ti} = \sum_{i=0}^{\infty} e^{-\lambda}\frac{\left(\lambda e^t\right)^i}{i!} = e^{-\lambda}e^{\lambda e^t} = e^{\lambda\left(e^t - 1\right)},$$

Here we used that $e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$. We have

$$
\begin{aligned}
M_K(t) = \mathbb{E}\left[e^{tK}\right] &= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{t\sum_{i=1}^{n} X_i}\right] \mathbb{P}(N=n) \\
&= \sum_{n=0}^{\infty} \left(q + pe^t\right)^n e^{-\lambda} \frac{\lambda^n}{n!} \\
&= \sum_{n=0}^{\infty} e^{-\lambda} \frac{\left(\lambda\left(q+pe^t\right)\right)^n}{n!} \\
&= e^{-\lambda} e^{\lambda\left(q+pe^t\right)} \\
&= e^{\lambda\left(q+pe^t-1\right)} \\
&= e^{\lambda\left(pe^t-p\right)} \\
&= e^{\lambda p\left(e^t-1\right)},
\end{aligned}
$$

we used the fact $q-1=-p$. This is the same as the MGF of $Poisson\,(\lambda p)$, thus $K \sim Poison\,(\lambda p)$.

4. **(Joint densities)** Let the joint density function of (X, Y) be

$$
f(x,y) = \begin{cases} 3xy(x+y), & \text{if } (x,y) \in [0,1]^2, \\ 0, & \text{else.} \end{cases}
$$

Calculate the covariance $\mathrm{Cov}(X,Y)$. We want to calculate $\mathbb{E}[X]$, $\mathbb{E}[Y]$ and $\mathbb{E}[XY]$. In all cases we need to apply the formula

$$
\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y)\, dx\, dy.
$$

Let's do each one of the calculations and then conclude.

$$
\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x,y)\, dx\, dy \\
&= \int_0^1 \int_0^1 3x^2 y(x+y)\, dx\, dy \\
&= \int_0^1 \int_0^1 3x^3 y + 3x^2 y^2\, dx\, dy \\
&= \int_0^1 \frac{3}{4} y + y^2\, dy \\
&= \frac{17}{24}.
\end{aligned}
$$

By symmetry we also have $\mathbb{E}[Y] = \dfrac{17}{24}$.

$$\begin{aligned}
\mathbb{E}[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x,y)\, dx\, dy \\
&= \int_0^1 \int_0^1 3x^2y^2(x+y)\, dx\, dy \\
&= \int_0^1 \int_0^1 3x^3y^2 + 3x^2y^3\, dx\, dy \\
&= \int_0^1 \frac{3}{4}y^2 + y^3\, dy \\
&= \frac{1}{2}.
\end{aligned}$$

We finally conclude using the formula $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 1/2 - (17/24)^2 = -1/576$.

5. **(Transformation of random variables)**

(a) Suppose X has the Cauchy distribution with density:

$$f_X(x) := \frac{1}{\pi(1+x^2)}.$$

Show that $1/X$ has the same distribution as X.
We aim to find the probability density function of $Y = \frac{1}{X}$. Using the change of variables formula, we have

$$\begin{aligned}
f_Y(y) &= f_X\left(\frac{1}{y}\right)\left|\frac{dx}{dy}\right| \\
&= \frac{1}{\pi\left(1+\left(\frac{1}{y}\right)^2\right)} \cdot \left|-\frac{1}{y^2}\right| \\
&= \frac{1}{\pi(1+y^2)}.
\end{aligned}$$

Therefore, $Y = \frac{1}{X}$ has the same distribution as X.

(b) Suppose $Y \sim \mathrm{Exp}(1)$. Find a function $g : (0,\infty) \to (-\infty,\infty)$ such that $g(Y)$ has the Cauchy distribution with density given by (a).
The CDF of X in (a) is

$$F_X(x) = \frac{1}{\pi}\left(\arctan(x) + \frac{\pi}{2}\right).$$

The CDF of Y is

$$F_Y(y) = P(Y \le y) = 1 - e^{-y}, \quad y > 0.$$

Suppose g is a strictly increasing function, we can relate the CDFs of Y and X as

$$F_Y(y) = P(Y \le y) = P(g(Y) \le g(y)) = F_X(g(y))$$

Substituting the CDFs, we have

$$1 - e^{-y} = \frac{1}{\pi}\left(\arctan(g(y)) + \frac{\pi}{2}\right)$$

Simplifying the above, we get

$$g(y) = \tan\left(\pi\left(1 - e^{-y}\right) - \frac{\pi}{2}\right)$$

which is strictly increasing when $y > 0$. This is the required transformation that ensures $g(Y)$ follows the Cauchy distribution.

(c) Suppose $Z \sim \text{Exp}(\lambda)$, where $\lambda > 0$. Show that the distribution of $W := \lceil Z \rceil$ (here $\lceil z \rceil$ is the smallest integer that is larger than or equal to z) is Geometric. Explicitly express the parameter of the Geometric distribution in terms of λ.

Suppose X has an exponential distribution with rate parameter λ, i.e., the density of X is given by

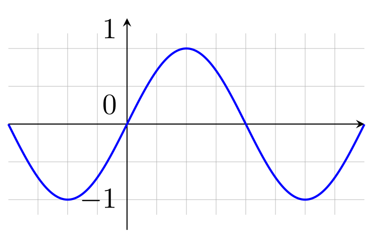
$$f_x(x) = \lambda e^{-\lambda x}; \quad x > 0.$$

Now consider $Y = \lceil X \rceil$. Since the ceiling function returns the smallest integer that is larger than or equal to X, Y is a discrete random variable. The PMF of Y can be expressed as

$$\begin{aligned}
\mathbb{P}[Y = y] &= \mathbb{P}[y - 1 < x \leq y] \\
&= \int_{y-1}^{y} \lambda e^{-\lambda x} dx \\
&= -e^{-\lambda x}\Big|_{y-1}^{y} \\
&= e^{-\lambda(y-1)}\left[1 - e^{-\lambda}\right] \\
&= \left(1 - \left(1 - e^{-\lambda}\right)\right)^{y-1}\left(1 - e^{-\lambda}\right)
\end{aligned}$$

The geometric distribution has the PMF $\mathbb{P}[X = x] = (1 - p)^{x-1}p$. Recognizing $p = 1 - e^{-\lambda}$, we now have the PMF of $Y = \lceil X \rceil$ in the form of a geometric distribution with parameter $p = 1 - e^{-\lambda}$. Therefore, for $X \sim \text{Exp}(\lambda)$, we have $Y = \lceil X \rceil \sim \text{Geo}(1 - e^{-\lambda})$.

6. **(Transformation of random variables)** Suppose $X \sim \text{Uniform}[-\pi, 2\pi]$. Find the p.d.f. of $Y = \sin(X)$.



Plot of $y = \sin(x)$.

Here we have a transformation of the form $Y = g(X)$ for $g(x) = \sin(x)$. While X takes values on $[-\pi, 2\pi]$, Y takes values on $[-1, 1]$. A picture of the function g on $[-\pi, 2\pi]$ show us that we have to make a distinction in two cases, either $y \in [-1, 0]$ or $y \in (0, 1]$. We want to use the formula $f_Y(y) = \sum_{\substack{g(x)=y,\\ g'(x)\neq 0}} f_X(x) \frac{1}{|g'(x)|}$. Note that $f_X(x) = 1/(3\pi)$ for $x \in [-\pi, 2\pi]$. $g'(x) = \cos(x)$. Now we separate in two cases.

(a) If $y \in (-1, 0)$ then y has 4 preimages, we obtain

$$\begin{aligned} f_Y(y) &= \sum_{\substack{g(x)=y,\\ g'(x)\neq 0}} f_X(x) \frac{1}{|g'(x)|} \\ &= \sum_{\substack{g(x)=y,\\ g'(x)\neq 0}} \frac{1}{3\pi} \times \frac{1}{\cos(\arcsin(y))} = \frac{4}{3\pi\sqrt{1-y^2}}. \end{aligned}$$

(b) If $y \in (1, 0)$ then y has 2 preimages, we obtain

$$\begin{aligned} f_Y(y) &= \sum_{\substack{g(x)=y,\\ g'(x)\neq 0}} f_X(x) \frac{1}{|g'(x)|} \\ &= \sum_{\substack{g(x)=y,\\ g'(x)\neq 0}} \frac{1}{3\pi} \times \frac{1}{\cos(\arcsin(y))} = \frac{2}{3\pi\sqrt{1-y^2}}. \end{aligned}$$

We shouldn't be concerned by the cases $y = -1, 0, 1$ since Y is a continuous random variable and we can modify the p.d.f at a finite number of points without any repercussion. We conclude that

$$f_Y(y) = \begin{cases} \frac{4}{3\pi\sqrt{1-y^2}} & \text{if } y \in (-1, 0), \\ \frac{2}{3\pi\sqrt{1-y^2}} & \text{if } y \in (0, 1), \\ 0 \text{ else.} \end{cases}$$

STAT201A_PS3_code

2024-10-06

Q1a

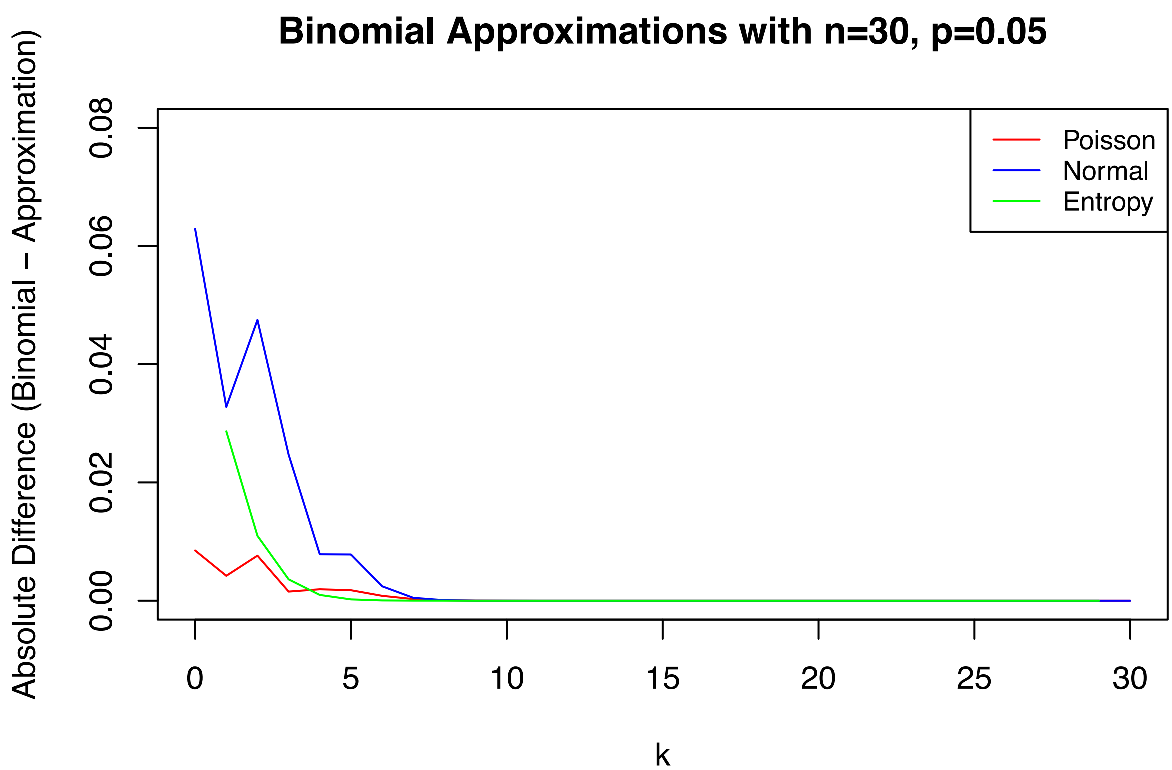
```r
## Given values
n <- 30
p <- 0.05
k <- 0:30
## Binomial Calculation
## P(Bin(n,p) = k)
binom <- dbinom(k, n, p)
## Poisson Approximation
## P(Pois(np) = k)
pois <- dpois(k, n*p)
## Normal Approximation
## phi(k; np, np(1-p))
norm <- dnorm(k, n*p, sqrt(n*p*(1-p)))
#n*abs(k/n - p)^3 ## normal approx is good if this value is small

## Entropy Approximation
## Ent(k; n, p)
## Note that f=k/n and the entropy approximation DNE for k=0 and k=30
f <- k/n
entr <- 1/(sqrt(2*pi*n*f*(1-f)))*exp(-n*(f*log(f/p) + (1-f)*log((1-f)/(1-p))))
## Error terms
## Binomial - Poisson
pois_diff <- abs(binom - pois)
## Binomial - Normal
norm_diff <- abs(binom - norm)
## Binomial - Entropy
entr_diff <- abs(binom - entr)
cbind(pois_diff, norm_diff, entr_diff)
```

```
##          pois_diff    norm_diff    entr_diff
##  [1,] 8.491396e-03 6.288532e-02          NaN
##  [2,] 4.208071e-03 3.277279e-02 2.865137e-02
##  [3,] 7.615308e-03 4.749378e-02 1.096979e-02
##  [4,] 1.538910e-03 2.470382e-02 3.605901e-03
##  [5,] 1.930467e-03 7.845186e-03 9.679277e-04
##  [6,] 1.766931e-03 7.810490e-03 2.143120e-04
##  [7,] 8.209923e-04 2.434696e-03 3.976033e-05
##  [8,] 2.675847e-04 4.806307e-04 6.268669e-06
##  [9,] 6.786047e-05 7.384760e-05 8.496632e-07
## [10,] 1.412178e-05 9.515639e-06 9.992828e-08
## [11,] 2.493920e-06 1.051824e-06 1.027277e-08
## [12,] 3.828577e-07 1.006534e-07 9.284106e-10
## [13,] 5.205110e-08 8.387780e-09 7.409012e-11
```

```
## [14,] 6.362462e-09 6.112552e-10 5.237982e-12
## [15,] 7.081188e-10 3.906518e-11 3.287950e-13
## [16,] 7.252526e-11 2.193133e-12 1.834906e-14
## [17,] 6.896636e-12 1.082138e-13 9.107894e-16
## [18,] 6.133846e-13 4.690382e-15 4.019292e-17
## [19,] 5.132796e-14 1.782894e-16 1.574848e-18
## [20,] 4.060356e-15 5.926516e-18 5.466524e-20
## [21,] 3.047997e-16 1.715570e-19 1.675528e-21
## [22,] 2.177936e-17 4.299675e-21 4.514869e-23
## [23,] 1.485157e-18 9.257674e-23 1.063399e-24
## [24,] 9.686240e-20 1.694769e-24 2.173290e-26
## [25,] 6.053980e-21 2.601619e-26 3.818433e-28
## [26,] 3.632400e-22 3.286255e-28 5.701307e-30
## [27,] 2.095617e-23 3.326169e-30 7.132860e-32
## [28,] 1.164232e-24 2.593504e-32 7.360840e-34
## [29,] 6.236956e-26 1.462502e-34 6.203044e-36
## [30,] 3.226012e-27 5.308539e-37 4.487914e-38
## [31,] 1.613006e-28 9.313226e-40          NaN
```

```r
plot(0:30, pois_diff, type="l", col="red", ylim=c(0,0.08),
ylab="Absolute Difference (Binomial - Approximation)", xlab="k",
main="Binomial Approximations with n=30, p=0.05") ### poisson best here since p is small
lines(0:30, norm_diff, col="blue")
lines(0:30, entr_diff, col="green")
legend("topright", legend=c("Poisson", "Normal", "Entropy"),
col=c("red", "blue", "green"), lty=c(1,1,1), cex=0.8)
```



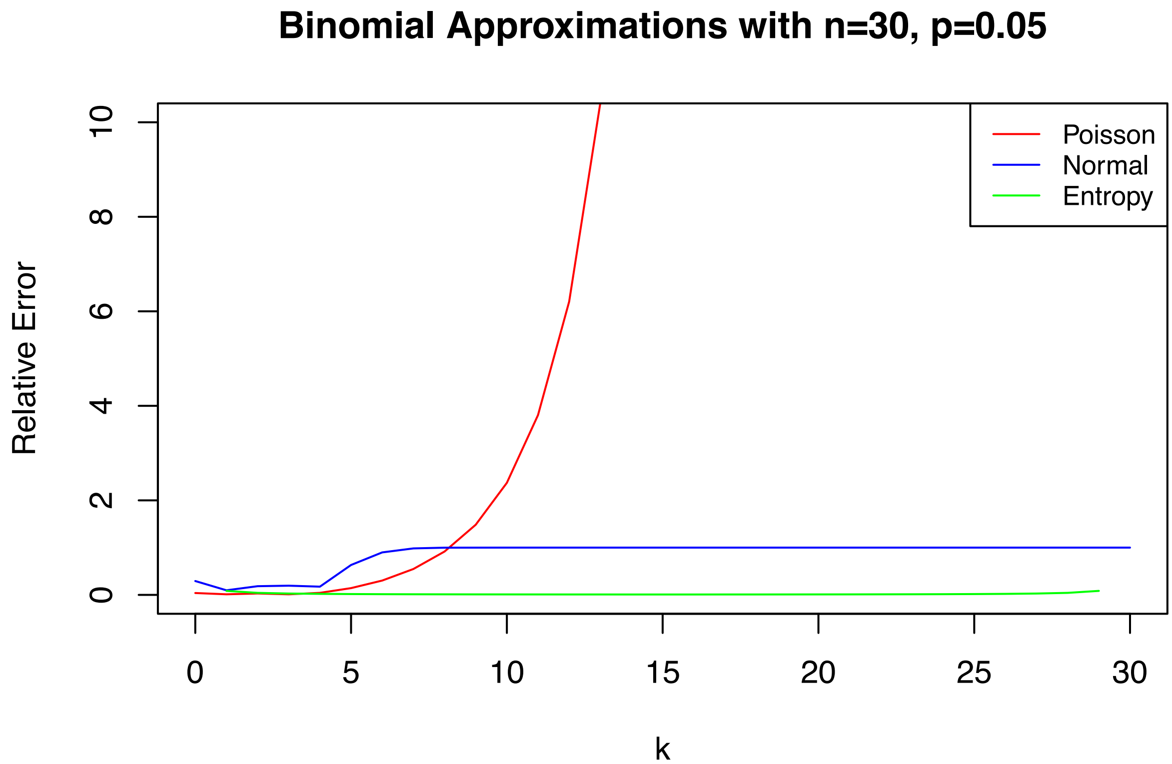
Based on this table, it seems that the Poisson approximation is doing a good job of approximating the Binomial distribution for small k. As k increases, we see that the error between the actual Binomial values and the approximations gets closer and closer to zero for all three approximations. It is important to note here though that the true Binomial values are inherently small, so the approximation differences are small. The Entropy approximation is not valid when $n-k$ or k is very small, which is why we have NaN values for $k=0$ and $k=n$.

Q1b

```
## Relative Errors
pois_rel_error <- pois_diff / binom
norm_rel_error <- norm_diff / binom
entr_rel_error <- entr_diff / binom
cbind(pois_rel_error, norm_rel_error, entr_rel_error)
```

```
##       pois_rel_error norm_rel_error entr_rel_error
##  [1,]   3.956134e-02     0.29298210            NaN
##  [2,]   1.241673e-02     0.09670249    0.084541418
##  [3,]   2.944403e-02     0.18363122    0.042413904
##  [4,]   1.211267e-02     0.19444231    0.028381831
##  [5,]   4.276996e-02     0.17381198    0.021444669
##  [6,]   1.430363e-01     0.63227346    0.017348947
##  [7,]   3.030614e-01     0.89874474    0.014677142
##  [8,]   5.473854e-01     0.98320354    0.012823520
##  [9,]   9.174123e-01     0.99835288    0.011486679
## [10,]   1.483921e+00     0.99990583    0.010500489
## [11,]   2.371035e+00     0.99999689    0.009766593
## [12,]   3.803725e+00     0.99999994    0.009223841
## [13,]   6.205587e+00     1.00000000    0.008833103
## [14,]   1.040885e+01     1.00000000    0.008569223
## [15,]   1.812660e+01     1.00000000    0.008416573
## [16,]   3.306925e+01     1.00000000    0.008366598
## [17,]   6.373157e+01     1.00000000    0.008416573
## [18,]   1.307750e+02     1.00000000    0.008569223
## [19,]   2.878913e+02     1.00000000    0.008833103
## [20,]   6.851169e+02     1.00000000    0.009223841
## [21,]   1.776666e+03     1.00000000    0.009766593
## [22,]   5.065349e+03     1.00000000    0.010500489
## [23,]   1.604244e+04     1.00000000    0.011486679
## [24,]   5.715375e+04     1.00000000    0.012823520
## [25,]   2.327005e+05     1.00000000    0.014677142
## [26,]   1.105331e+06     1.00000000    0.017348947
## [27,]   6.300392e+06     1.00000000    0.021444669
## [28,]   4.489030e+07     1.00000000    0.028381831
## [29,]   4.264579e+08     1.00000000    0.042413904
## [30,]   6.077024e+09     1.00000000    0.084541418
## [31,]   1.731952e+11     1.00000000            NaN
```

```
plot(0:30, pois_rel_error, type="l", col="red", ylim=c(0,10),
ylab="Relative Error", xlab="k",
main="Binomial Approximations with n=30, p=0.05") ### poisson best here since p is small
lines(0:30, norm_rel_error, col="blue")
lines(0:30, entr_rel_error, col="green")
legend("topright", legend=c("Poisson", "Normal", "Entropy"),
col=c("red", "blue", "green"), lty=c(1,1,1), cex=0.8)
```



The relative error is pretty small for all three approximations until we hit about $k = 5$. Overall, the relative Entropy error is best and is consistently small. The Poisson approximation holds when we look at absolute difference, but it does not hold in the context of relative error, which is why see a spike in relative error for Poisson. For the Normal relative error, it stays at about 1 for $k = 8$ and beyond.

Q1c

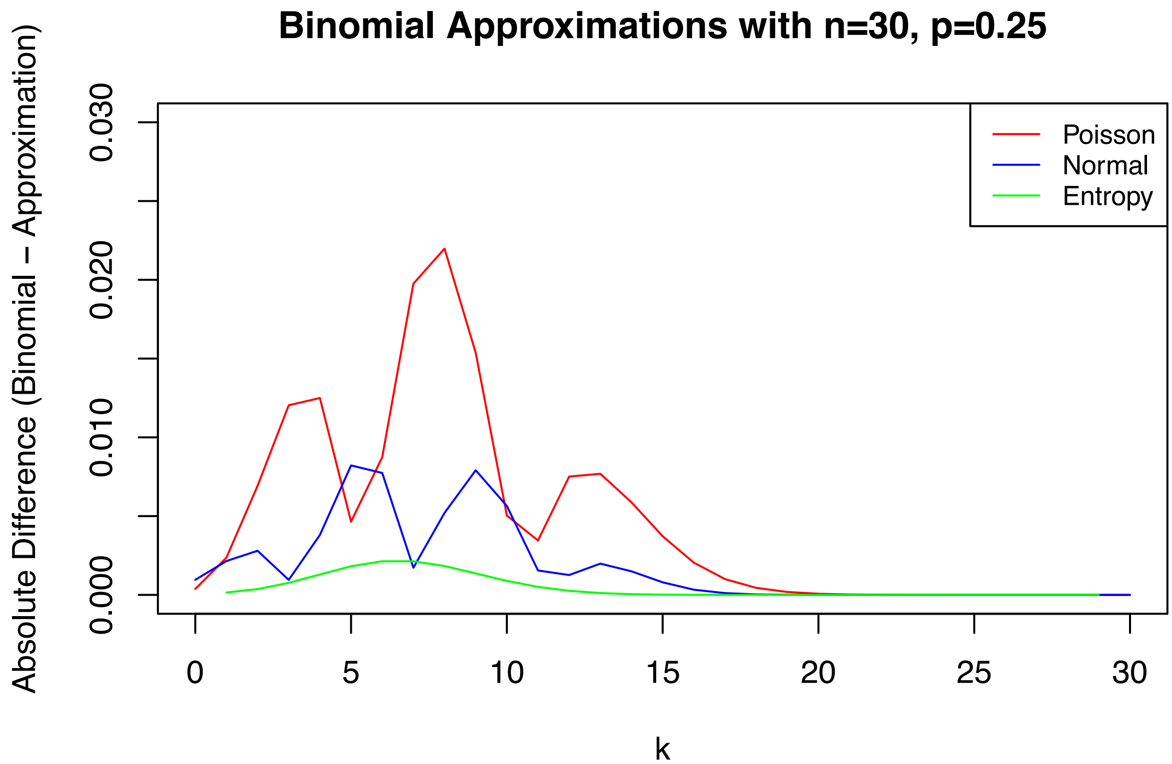
```
## Repeat this process for n = 30 and p = 0.25
n <- 30
p <- 0.25
k <- 0:30
## Binomial Calculation
## P(Bin(n,p) = k)
binom2 <- dbinom(k, n, p)
## Poisson Approximation
## P(Pois(np) = k)
pois2 <- dpois(k, n*p)
## Normal Approximation
## phi(k; np, np(1-p))
norm2 <- dnorm(k, n*p, sqrt(n*p*(1-p)))
#n*abs(k/n - p)^3 ## normal approx is good if this value is small

## Entropy Approximation
## Ent(k; n, p)
## Note that f=k/n and the entropy approximation DNE for k=0 and k=30
f2 <- k/n
entr2 <- 1/(sqrt(2*pi*n*f*(1-f)))*exp(-n*(f*log(f/p) + (1-f)*log((1-f)/(1-p))))
## Error terms
## Binomial - Poisson
```

```r
pois_diff2 <- abs(binom2 - pois2)
## Binomial - Normal
norm_diff2 <- abs(binom2 - norm2)
## Binomial - Entropy
entr_diff2 <- abs(binom2 - entr2)
cbind(pois_diff2, norm_diff2, entr_diff2)
```

```
##         pois_diff2   norm_diff2   entr_diff2
##  [1,] 3.745023e-04 9.548001e-04          NaN
##  [2,] 2.362312e-03 2.148155e-03 1.509758e-04
##  [3,] 6.924030e-03 2.799400e-03 3.660942e-04
##  [4,] 1.203529e-02 9.512783e-04 7.621502e-04
##  [5,] 1.249612e-02 3.802816e-03 1.295693e-03
##  [6,] 4.646120e-03 8.218080e-03 1.816929e-03
##  [7,] 8.737971e-03 7.738469e-03 2.134882e-03
##  [8,] 1.975184e-02 1.723586e-03 2.131727e-03
##  [9,] 2.198059e-02 5.202900e-03 1.829934e-03
## [10,] 1.536699e-02 7.910260e-03 1.363042e-03
## [11,] 5.034870e-03 5.645154e-03 8.874438e-04
## [12,] 3.450864e-03 1.547615e-03 5.079555e-04
## [13,] 7.510803e-03 1.259906e-03 2.567309e-04
## [14,] 7.686768e-03 1.983581e-03 1.149514e-04
## [15,] 5.874565e-03 1.495681e-03 4.569911e-05
## [16,] 3.721567e-03 7.971628e-04 1.615209e-05
## [17,] 2.046132e-03 3.299491e-04 5.077679e-06
## [18,] 1.003255e-03 1.104228e-04 1.419153e-06
## [19,] 4.471579e-04 3.054174e-05 3.521686e-07
## [20,] 1.838540e-04 7.073812e-06 7.742043e-08
## [21,] 7.055401e-05 1.382503e-06 1.502894e-08
## [22,] 2.550318e-05 2.287576e-07 2.564804e-09
## [23,] 8.744228e-06 3.202122e-08 3.825936e-10
## [24,] 2.858378e-06 3.772368e-09 4.952123e-11
## [25,] 8.940745e-07 3.702489e-10 5.510502e-12
## [26,] 2.683049e-07 2.978268e-11 5.210901e-13
## [27,] 7.740240e-08 1.915055e-12 4.128901e-14
## [28,] 2.150111e-08 9.472807e-14 2.698550e-15
## [29,] 5.759247e-09 3.385663e-15 1.440258e-16
## [30,] 1.489461e-09 7.782203e-17 6.599519e-18
## [31,] 3.723653e-10 8.625467e-19          NaN
```

```r
plot(0:30, pois_diff2, type="l", col="red", ylim=c(0,0.03),
ylab="Absolute Difference (Binomial - Approximation)", xlab="k",
main="Binomial Approximations with n=30, p=0.25")
lines(0:30, norm_diff2, col="blue")
lines(0:30, entr_diff2, col="green") ## entropy is best here
legend("topright", legend=c("Poisson", "Normal", "Entropy"),
col=c("red", "blue", "green"), lty=c(1,1,1), cex=0.8)
```

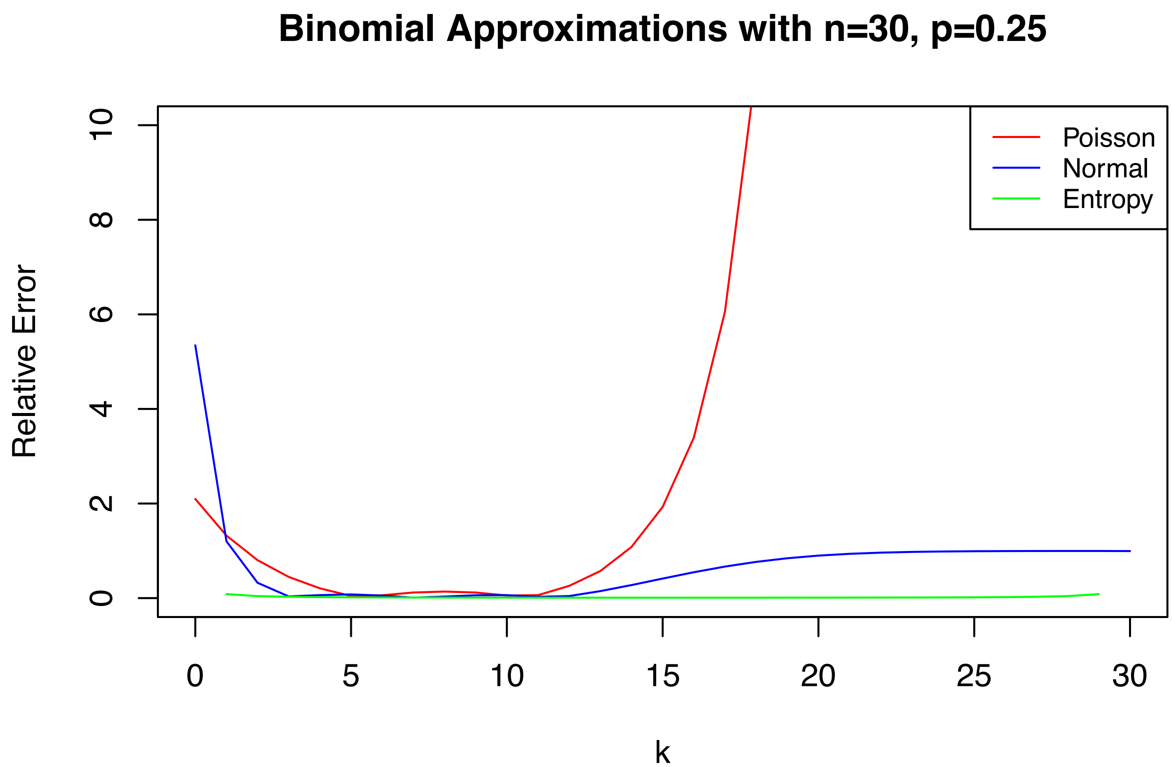



```
## Relative Errors
pois_rel_error2 <- pois_diff2 / binom2
norm_rel_error2 <- norm_diff2 / binom2
entr_rel_error2 <- entr_diff2 / binom2
cbind(pois_rel_error2, norm_rel_error2, entr_rel_error2)
```

```
##       pois_rel_error2 norm_rel_error2 entr_rel_error2
##  [1,]    2.097088e+00      5.34656148             NaN
##  [2,]    1.322816e+00      1.20289525     0.084541418
##  [3,]    8.021846e-01      0.32432490     0.042413904
##  [4,]    4.481840e-01      0.03542480     0.028381831
##  [5,]    2.068200e-01      0.06293940     0.021444669
##  [6,]    4.436348e-02      0.07847035     0.017348947
##  [7,]    6.007286e-02      0.05320136     0.014677142
##  [8,]    1.188183e-01      0.01036833     0.012823520
##  [9,]    1.379744e-01      0.03265913     0.011486679
## [10,]    1.183829e-01      0.06093839     0.010500489
## [11,]    5.541030e-02      0.06212666     0.009766593
## [12,]    6.266342e-02      0.02810277     0.009223841
## [13,]    2.584172e-01      0.04334841     0.008833103
## [14,]    5.730215e-01      0.14786901     0.008569223
## [15,]    1.081940e+00      0.27546514     0.008416573
## [16,]    1.927728e+00      0.41292112     0.008366598
## [17,]    3.391593e+00      0.54691136     0.008416573
## [18,]    6.057917e+00      0.66676229     0.008569223
## [19,]    1.121563e+01      0.76604869     0.008833103
## [20,]    2.190430e+01      0.84277120     0.009223841
## [21,]    4.584970e+01      0.89842291     0.009766593
## [22,]    1.044118e+02      0.93654972     0.010500489
## [23,]    2.625296e+02      0.96137894     0.011486679
```

```
## [24,]    7.401769e+02      0.97685460     0.012823520
## [25,]    2.381354e+03      0.98615246     0.014677142
## [26,]    8.932828e+03      0.99157154     0.017348947
## [27,]    4.020123e+04      0.99464042     0.021444669
## [28,]    2.261365e+05      0.99629656     0.028381831
## [29,]    1.696030e+06      0.99703786     0.042413904
## [30,]    1.908035e+07      0.99691874     0.084541418
## [31,]    4.293080e+08      0.99444867             NaN
```

```r
plot(0:30, pois_rel_error2, type="l", col="red", ylim=c(0,10),
ylab="Relative Error", xlab="k",
main="Binomial Approximations with n=30, p=0.25") ### poisson best here since p is small
lines(0:30, norm_rel_error2, col="blue")
lines(0:30, entr_rel_error2, col="green")
legend("topright", legend=c("Poisson", "Normal", "Entropy"),
col=c("red", "blue", "green"), lty=c(1,1,1), cex=0.8)
```



As shown in the above plot and tables, the entropy approximation is behaving best here with $p = 0.25$. The absolute difference for entropy is consistently small and the overall relative errors are smaller for entropy when compared to the other two approximations. At $k = 7$, we actually do see that the Normal approximation has a smaller absolute difference than the Entropy approximation, but overall the Entropy approximations behave the best. In terms of relative error, for small k Normal has the worst relative error. For k greater than 14, the Poisson relative error is the worst. The relative error for Entropy is consistently small.

Q1d

```r
## Repeat this process for n = 30 and p = 0.5
n <- 30
p <- 0.5
k <- 0:30
```

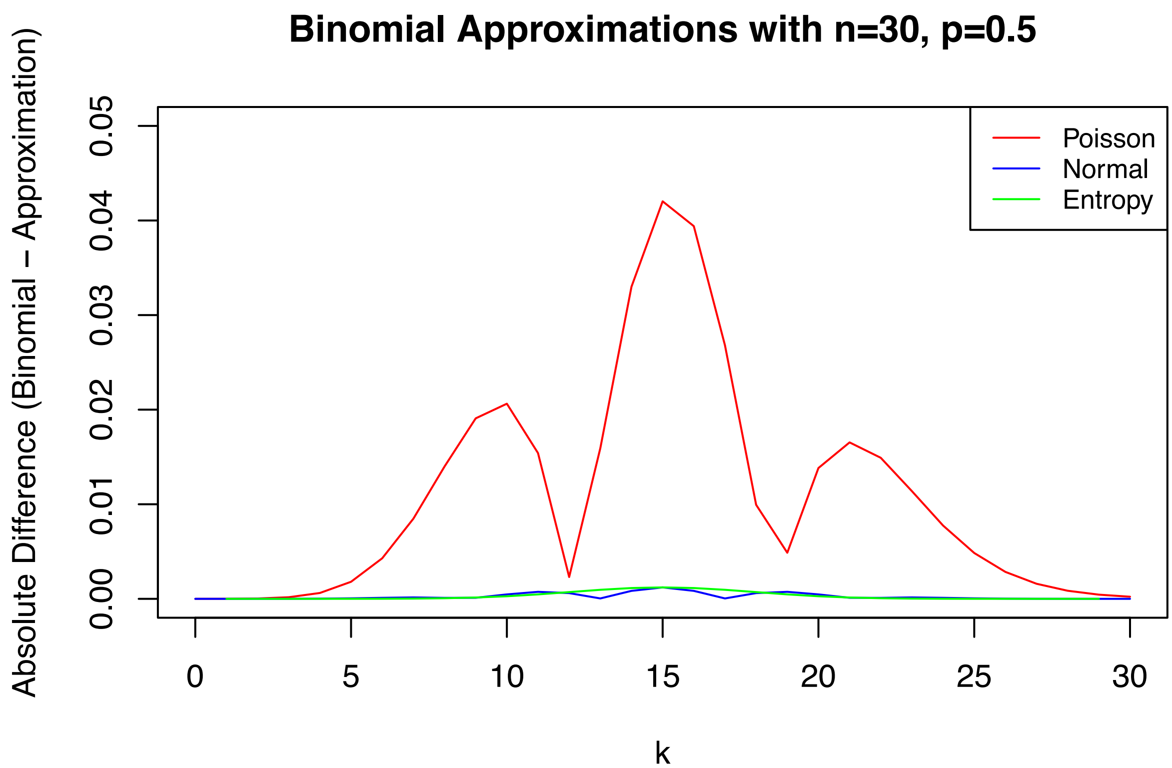
```r
## Binomial Calculation
## P(Bin(n,p) = k)
binom3 <- dbinom(k, n, p)
## Poisson Approximation
## P(Pois(np) = k)
pois3 <- dpois(k, n*p)
## Normal Approximation
norm3 <- dnorm(k, n*p, sqrt(n*p*(1-p)))
#n*abs(k/n - p)^3 ## normal approx is good if this value is small

## Entropy Approximation
## Ent(k; n, p)
## Note that f=k/n and the entropy approximation DNE for k=0 and k=30
f3 <- k/n
entr3 <- 1/(sqrt(2*pi*n*f*(1-f)))*exp(-n*(f*log(f/p) + (1-f)*log((1-f)/(1-p))))
## Error terms
## Binomial - Poisson
pois_diff3 <- abs(binom3 - pois3)
## Binomial - Normal
norm_diff3 <- abs(binom3 - norm3)
## Binomial - Entropy
entr_diff3 <- abs(binom3 - entr3)
cbind(pois_diff3, norm_diff3, entr_diff3)
```

```
##         pois_diff3   norm_diff3   entr_diff3
##   [1,] 3.049710e-07 4.363042e-08          NaN
##   [2,] 4.560595e-06 2.800940e-07 2.362060e-09
##   [3,] 3.400889e-05 1.458370e-06 1.718295e-08
##   [4,] 1.682889e-04 6.085087e-06 1.073165e-07
##   [5,] 6.197398e-04 2.019336e-05 5.473300e-07
##   [6,] 1.803069e-03 5.266949e-05 2.302536e-06
##   [7,] 4.286474e-03 1.049484e-04 8.116402e-06
##   [8,] 8.474307e-03 1.475841e-04 2.431322e-05
##   [9,] 1.399334e-02 1.040396e-04 6.261344e-05
##  [10,] 1.908260e-02 1.094041e-04 1.399145e-04
##  [11,] 2.062915e-02 4.675016e-04 2.732849e-04
##  [12,] 1.541175e-02 7.416804e-04 4.692688e-04
##  [13,] 2.306141e-03 6.059874e-04 7.115337e-04
##  [14,] 1.592824e-02 4.012203e-05 9.557687e-04
##  [15,] 3.299955e-02 8.428050e-04 1.139902e-03
##  [16,] 4.202858e-02 1.208676e-03 1.208676e-03
##  [17,] 3.940180e-02 8.428050e-04 1.139902e-03
##  [18,] 2.679950e-02 4.012203e-05 9.557687e-04
##  [19,] 9.940133e-03 6.059874e-04 7.115337e-04
##  [20,] 4.871436e-03 7.416804e-04 4.692688e-04
##  [21,] 1.382870e-02 4.675016e-04 2.732849e-04
##  [22,] 1.653993e-02 1.094041e-04 1.399145e-04
##  [23,] 1.491120e-02 1.040396e-04 6.261344e-05
##  [24,] 1.138368e-02 1.475841e-04 2.431322e-05
##  [25,] 7.746798e-03 1.049484e-04 8.116402e-06
##  [26,] 4.847157e-03 5.266949e-05 2.302536e-06
##  [27,] 2.847483e-03 2.019336e-05 5.473300e-07
##  [28,] 1.592333e-03 6.085087e-06 1.073165e-07
##  [29,] 8.546561e-04 1.458370e-06 1.718295e-08
```

```
## [30,] 4.422451e-04 2.800940e-07 2.362060e-09
## [31,] 2.211356e-04 4.363042e-08          NaN
```

```
plot(0:30, pois_diff3, type="l", col="red", ylim=c(0,0.05),
ylab="Absolute Difference (Binomial - Approximation)", xlab="k",
main="Binomial Approximations with n=30, p=0.5") ### poisson best here since p is small
lines(0:30, norm_diff3, col="blue")
lines(0:30, entr_diff3, col="green")
legend("topright", legend=c("Poisson", "Normal", "Entropy"),
col=c("red", "blue", "green"), lty=c(1,1,1), cex=0.8)
```

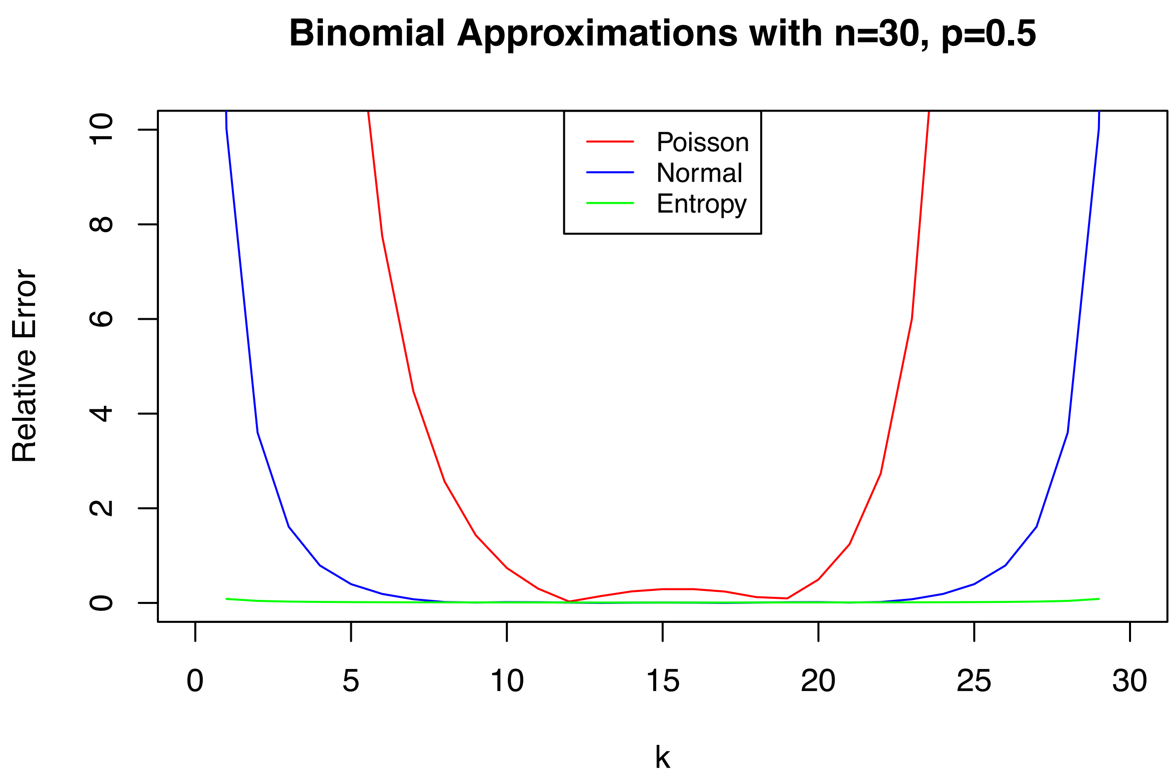



```
pois_rel_error3 <- pois_diff3 / binom3
norm_rel_error3 <- norm_diff3 / binom3
entr_rel_error3 <- entr_diff3 / binom3
cbind(pois_rel_error3, norm_rel_error3, entr_rel_error3)
```

```
##       pois_rel_error3 norm_rel_error3 entr_rel_error3
##  [1,]    3.274601e+02    4.684781e+01             NaN
##  [2,]    1.632301e+02    1.002495e+01     0.084541418
##  [3,]    8.394658e+01    3.599799e+00     0.042413904
##  [4,]    4.450710e+01    1.609313e+00     0.028381831
##  [5,]    2.428172e+01    7.911860e-01     0.021444669
##  [6,]    1.358561e+01    3.968495e-01     0.017348947
##  [7,]    7.751365e+00    1.897814e-01     0.014677142
##  [8,]    4.469603e+00    7.784027e-02     0.012823520
##  [9,]    2.567132e+00    1.908647e-02     0.011486679
## [10,]    1.432136e+00    8.210703e-03     0.010500489
## [11,]    7.372398e-01    1.670746e-02     0.009766593
## [12,]    3.029299e-01    1.457830e-02     0.009223841
## [13,]    2.862884e-02    7.522833e-03     0.008833103
## [14,]    1.428093e-01    3.597258e-04     0.008569223
```

```
## [15,]     2.436553e-01    6.222929e-03     0.008416573
## [16,]     2.909268e-01    8.366598e-03     0.008366598
## [17,]     2.909268e-01    6.222929e-03     0.008416573
## [18,]     2.402787e-01    3.597258e-04     0.008569223
## [19,]     1.233985e-01    7.522833e-03     0.008833103
## [20,]     9.575183e-02    1.457830e-02     0.009223841
## [21,]     4.942070e-01    1.670746e-02     0.009766593
## [22,]     1.241311e+00    8.210703e-03     0.010500489
## [23,]     2.735518e+00    1.908647e-02     0.011486679
## [24,]     6.004096e+00    7.784027e-02     0.012823520
## [25,]     1.400878e+01    1.897814e-01     0.014677142
## [26,]     3.652194e+01    3.968495e-01     0.017348947
## [27,]     1.115658e+02    7.911860e-01     0.021444669
## [28,]     4.211218e+02    1.609313e+00     0.028381831
## [29,]     2.109609e+03    3.599799e+00     0.042413904
## [30,]     1.582857e+04    1.002495e+01     0.084541418
## [31,]     2.374425e+05    4.684781e+01             NaN
```

```
plot(0:30, pois_rel_error3, type="l", col="red", ylim=c(0,10),
ylab="Relative Error", xlab="k",
main="Binomial Approximations with n=30, p=0.5") ### poisson best here since p is small
lines(0:30, norm_rel_error3, col="blue")
lines(0:30, entr_rel_error3, col="green")
legend("top", legend=c("Poisson", "Normal", "Entropy"),
col=c("red", "blue", "green"), lty=c(1,1,1), cex=0.8)
```



The absolute error for Poisson is the worst among our approximations. The Poisson absolute errors have a multi-modal shape, indicating that the Poisson approximations are only suitable for select values of k. The Normal and Entropy approximations are small overall. Within the range of $k = 12, 13, 14, 16, 17, 18$ we see that the Normal approximation is actually lower than Entropy in terms of absolute error. In this scenario, Normal is the best approximation. For relative errors, Entropy again is the best. The Entropy relative errors are consistently small. From $k = 7$ to $k = 23$, the Normal relative errors are pretty small. Outside of this range, the approximation isn't appropriate. For $k = 12$ and $k = 19$, the Poisson relative error is small, but outside of this range, the approximation worsens.