

Question 1

A *probability mass function* operates on discrete random variables. Its evaluation returns the probability that a random variable will take on the value of its argument. A *probability density function*, on the other hand, operates over continuous random variables. Probability density functions do not return probability values, but probability densities. Integration must be applied to return the probability over an interval.

As stated in the class textbook, a *random variable* is a quantity that can take on values due to chance. They do not have a single value, but instead can take on a range of values. The important distinction between random variables and the parameters of a probability distribution is that random variables are governed by probability distributions (and hence the parameters of the distributions). In some cases, the parameters of a distribution can also be modeled as random variables.

Question 2

The following plots demonstrate convergence of the Binomial and Poisson distributions to a Normal. It can be observed in Figure 1, that as the λ parameter of the Poisson grows to infinity, the distribution of samples converges to a Normal centered at λ . The same is true for the Binomial in Figure 2. For a fixed probability of success, $p = 0.5$, the distribution converges to a Normal.

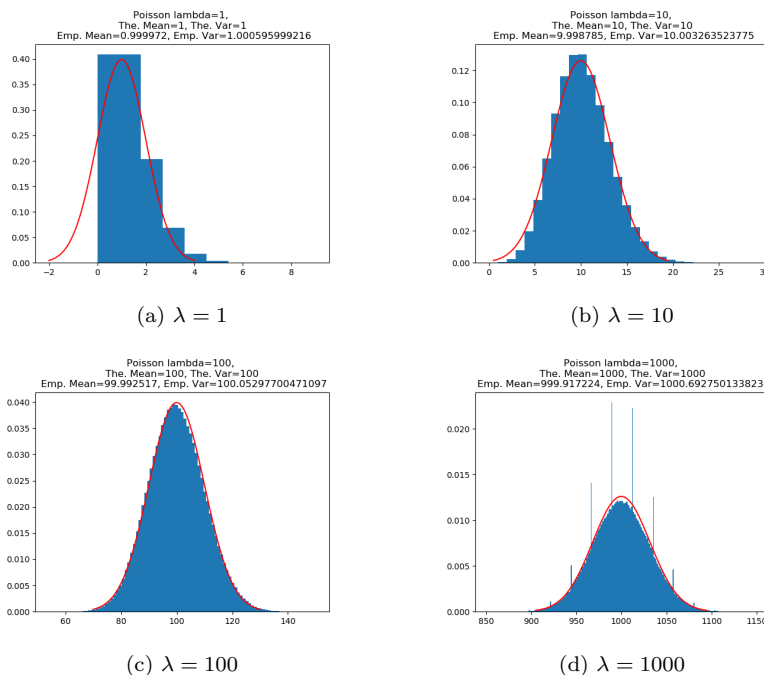
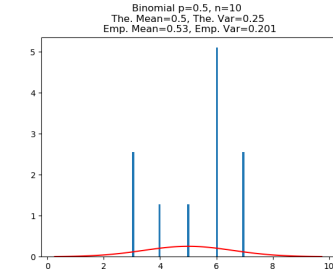
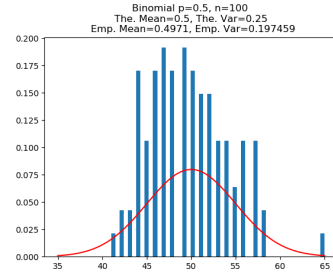


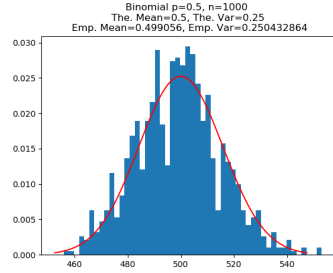
Figure 1: Convergence of the Poisson distribution to a Normal distribution for a fixed number of samples, $n = 100000$. As the mean parameter λ goes to infinity, the distribution of Poisson samples converges to a Normal.



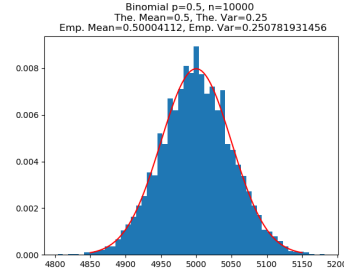
(a) $n = 10$



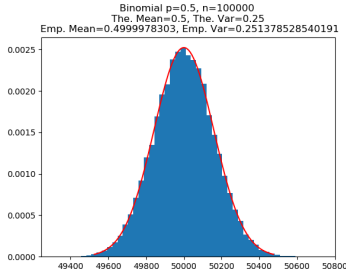
(b) $n = 100$



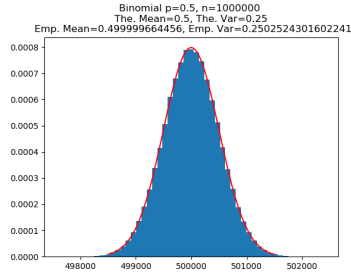
(c) $n = 1000$



(d) $n = 10000$



(e) $n = 100000$



(f) $n = 1000000$

Figure 2: Convergence of the Binomial distribution to a Normal distribution. The true p value of the Binomial was fixed at $p = 0.5$. As N goes to infinity, the distribution of Binomial samples converges to a Normal.

Question 3

For the Gamma distribution, $\text{gamma}(\alpha, \beta)$, it is known that the skewness is parameterized by the shape, or $\gamma = \frac{2}{\sqrt{\alpha}}$. The class text book used moment matching with the Gamma to find the mean and variance as $\mu = \frac{\alpha}{\beta}$ and $\sigma^2 = \frac{\alpha}{\beta^2}$, respectively. Using moment matching, we can find the skewness as $\gamma = \frac{\mu_3}{(\sigma^2)^{\frac{3}{2}}}$, where $\mu_3 = \mathbb{E}[(X - \mu)^3] = \mathbb{E}[X^3] - 3\mu\sigma^2 - \mu^3$ and $\mathbb{E}[X^3] = \frac{(\alpha+2)(\alpha+1)\alpha}{\beta^3}$. Therefore,

$$\begin{aligned}\gamma &= \frac{\mu_3}{(\sigma^2)^{\frac{3}{2}}} \\ &= \frac{\mathbb{E}[X^3] - 3\mu\sigma^2 - \mu^3}{(\sigma^2)^{\frac{3}{2}}} \\ &= \frac{\frac{(\alpha+2)(\alpha+1)\alpha}{\beta^3} - 3\left(\frac{\alpha}{\beta}\right)\left(\frac{\alpha}{\beta^2}\right) - \left(\frac{\alpha}{\beta}\right)^3}{\left(\frac{\alpha}{\beta^2}\right)^{\frac{3}{2}}} \\ &= \frac{2}{\sqrt{\alpha}}\end{aligned}$$

Figure 3 demonstrates pdfs of the Gamma distribution for various shape parameters, α .

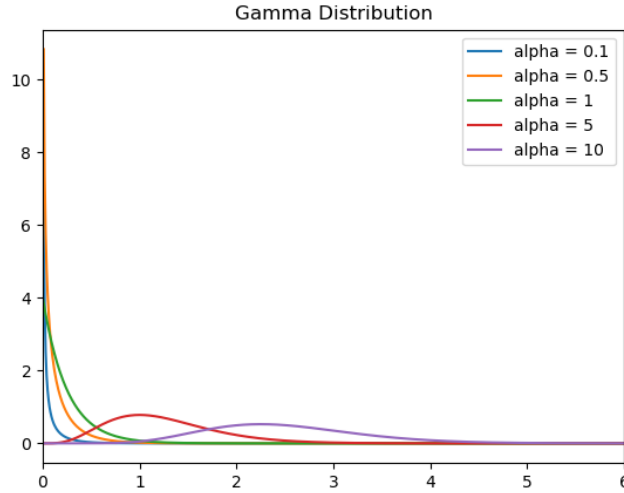


Figure 3: Gamma pdf for various shape parameters, α .

Question 4

Binomial sampling experiments were repeated 10^5 times for the parameterizations requested in the assignment handout. The standard errors for were computed for each trial. The results are shown in Table 1. As can be observed, the deviation of the error decreases as the number of flips increases.

Table 1: Standard deviation of the point error for the proportion of heads from Binomial samples.

Number of Flips	5	10	50	100
Standard Error	0.205	0.145	0.065	0.046

Manipulating the Chebychev inequality, we can place bounds on the expected value of the proportion of heads. The Chebychev inequality states that for a random variable X whose governing distribution has mean μ , that

$$P\{|X - \mu| \geq d\} \leq \frac{\sigma^2}{d^2}$$

Then for the standard error of the estimate for p , denoted as \bar{p} :

$$P\left\{|\bar{p} - p| \leq 2\frac{\sigma}{\sqrt{N}}\right\} \approx 0.95$$

Question 5

The maximum likelihood estimator for the p value of a Binomial was found as:

$$\hat{p} = \frac{\sum_{i=1}^N y_i}{Nn}$$

The derivation is as follows:

$$\begin{aligned} [Y|n, p] &= \prod_{i=1}^N \binom{n}{y_i} p^{y_i} (1-p)^{(n-y_i)} \\ \Rightarrow \log[Y|n, p] &= \log \prod_{i=1}^N \binom{n}{y_i} p^{y_i} (1-p)^{(n-y_i)} \\ &= \sum_{i=1}^N \left[\log \binom{n}{y_i} + y_i \log p + (n - y_i) \log(1-p) \right] \\ &= \sum_{i=1}^N \log \binom{n}{y_i} + \log p \sum_{i=1}^N y_i + \log(1-p) \sum_{i=1}^N (n - y_i) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial}{\partial p} [Y|n, p] &= \frac{\partial}{\partial p} \left[\sum_{i=1}^N \log \binom{n}{y_i} + \log p \sum_{i=1}^N y_i + \log(1-p) \sum_{i=1}^N (n - y_i) \right] \\ &= \frac{1}{p} \sum_{i=1}^N y_i - \frac{1}{1-p} \sum_{i=1}^N (n - y_i) \end{aligned}$$

Equating to zero provides

$$\begin{aligned} \frac{1}{p} \sum_{i=1}^N y_i - \frac{1}{1-p} \sum_{i=1}^N (n - y_i) &= 0 \\ \Rightarrow (1-p) \sum_{i=1}^N y_i &= p \sum_{i=1}^N (n - y_i) \\ \Rightarrow p &= \frac{\sum_{i=1}^N y_i}{\sum_{i=1}^N n} = \frac{\sum_{i=1}^N y_i}{Nn} \end{aligned}$$

The maximum likelihood estimator was applied to each of the datasets for p . Figure 4 demonstrates the histograms of the samples for repeated draws from Binomials. The MLE estimates for p as well as the number of flips in each Binomial trial are shown in the titles.

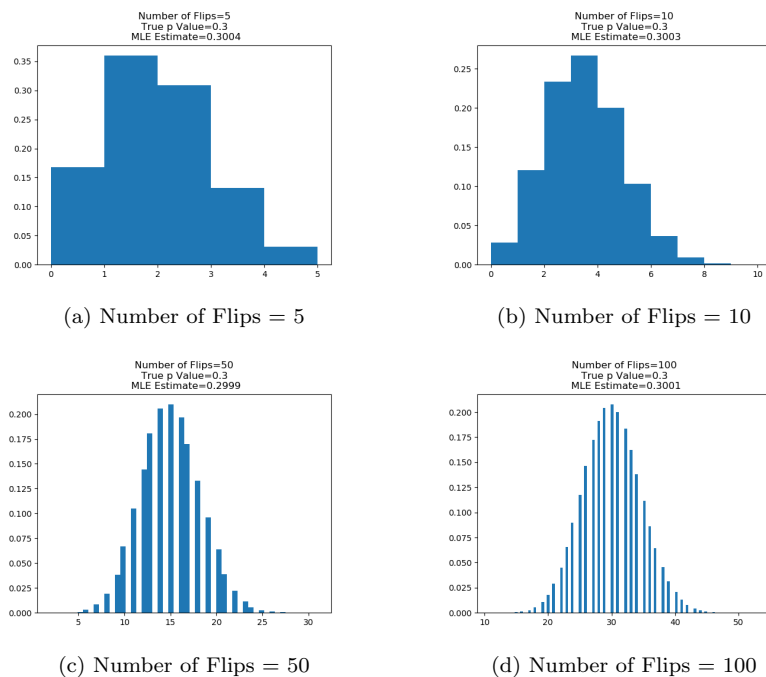


Figure 4: Histograms of samples drawn from Binomial distributions parameterized by various numbers of flips.

Figure 4 demonstrates the standard errors from 100 Monte Carlo simulations for various parameterizations of the Binomial distribution. As seen in the figure, the standard deviation of the parameter estimate error decreases as the number of flips increases.

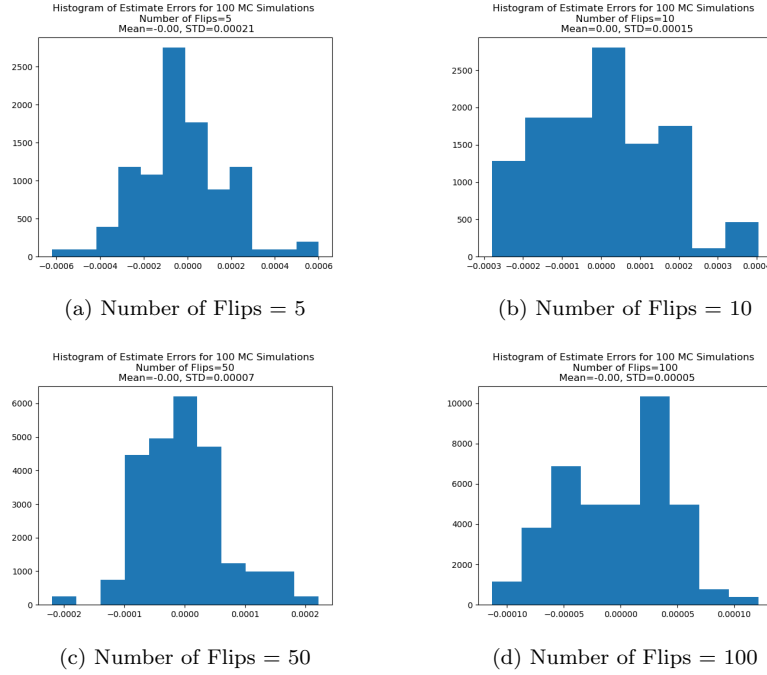


Figure 5: Histograms of the parameter estimate error for p . Each histogram represents the estimate errors

Question 6

Question 6 provided us the differential equation:

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right)$$

By observation, $\lim_{N \rightarrow \infty} \frac{dN}{dt} = -\infty$. The function attains its maximum value at an inflection point, found as

$$\begin{aligned} \frac{\partial}{\partial N} \frac{dN}{dt} &= \frac{\partial}{\partial N} rN - \frac{rN^2}{K} \\ &= r - \frac{2rN}{K} \\ &= 0 \\ \Rightarrow N &= \frac{K}{2} \end{aligned}$$

This is reinforced by Figure 6, which shows trajectories of $\frac{dN}{dt}$ for various values of r and K .

The process was then implemented as described in the homework handout. Figure 7 demonstrates 10^3 sample paths of the process. Zero-mean observation error parameterized by a $\text{Normal}(0, 0.5)$ was applied to the measurements. It can be observed that, while the paths showed deviation, they were relatively stable with a consistent mean.

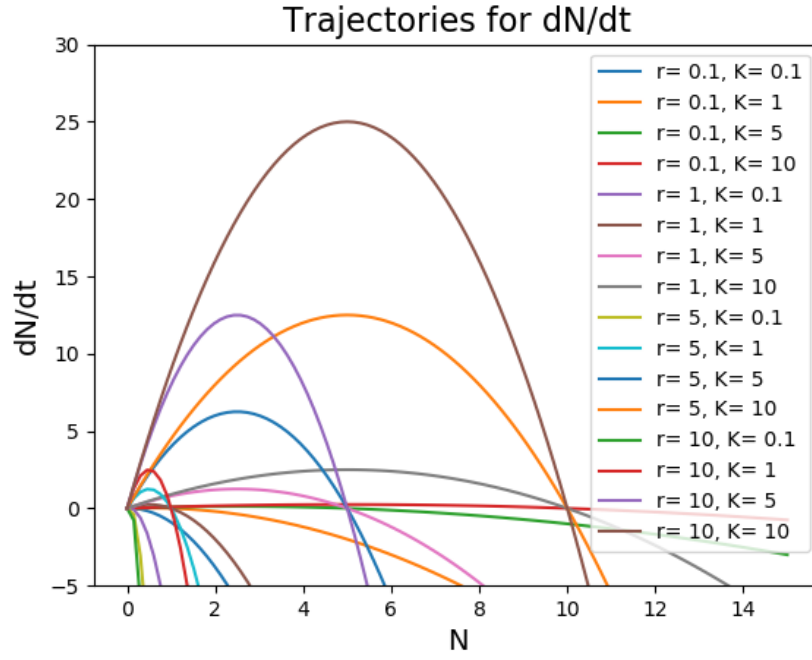


Figure 6: Trajectories of $\frac{dN}{dt}$ for various values of r and K . Each trajectory attain its maximum at $N = \frac{K}{2}$.

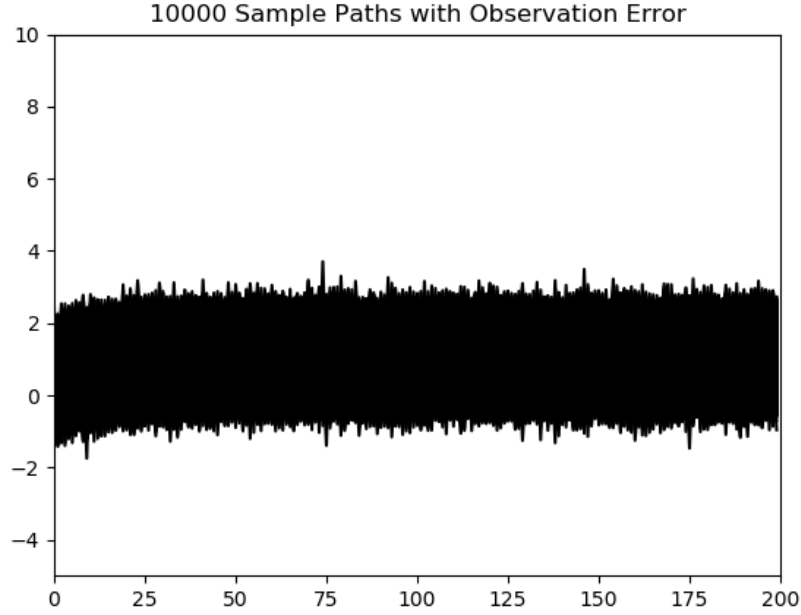


Figure 7: Sample paths of the process N with zero-mean observation error.

The previous trial was repeated with process error instead of observation error. Figure 8 shows the resulting 10^3 sample paths.

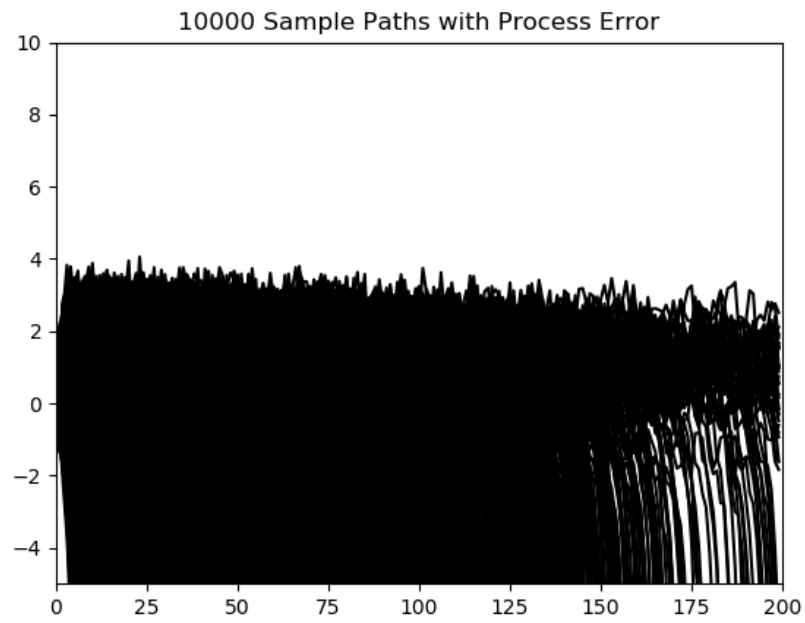


Figure 8

It can be observed that the process error avalanches for specific sample paths, thus forcing the trajectories to $-\infty$.