1.

$$\widehat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2$$

$$\langle\widehat{\sigma^2}\rangle = \frac{1}{n}\sum_{i=1}^{n}\langle(x_i-\hat{\mu})^2\rangle = \frac{1}{n}\sum_{i=1}^{n}\langle x_i^2\rangle - 2\langle x_i\hat{\mu}\rangle + \langle\hat{\mu}^2\rangle$$

$$[\hat{\mu} = \bar{\mu}]$$

$$\langle x_i^2\rangle = \int_{-\infty}^{\infty}\frac{x_i^2}{\sqrt{2\pi}\,\sigma}e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}dx_i$$

from a table, we find

$$\int_{-\infty}^{\infty}\frac{x^2}{\sqrt{2\pi}\,\sigma}e^{\frac{-(x-\mu)^2}{2\sigma^2}}dx = \sigma^2+\mu^2$$

$$\Rightarrow \langle x_i^2\rangle = \sigma^2+\mu^2$$

$$\langle x_i\hat{\mu}\rangle = \langle\frac{1}{n}\sum_{j=1}^{n}x_ix_j\rangle$$

$$= \frac{1}{n}\Big(\langle x_i^2\rangle + \sum_{\substack{j=1\\ j\neq i}}^{n}\langle x_ix_j\rangle\Big)$$

$$\langle x_ix_j\rangle = \int_{-\infty}^{\infty}dx_i\int_{-\infty}^{\infty}dx_j\,(x_ix_j)\frac{1}{2\pi\sigma^2}e^{\frac{-(x_i-\mu)^2}{2\sigma^2}-\frac{(x_j-\mu)^2}{2\sigma^2}}$$

$$= \int_{-\infty}^{\infty}dx_i\,\frac{x_i}{\sqrt{2\pi}\,\sigma}e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}\int_{-\infty}^{\infty}\frac{x_j}{\sqrt{2\pi}\,\sigma}e^{\frac{-(x_j-\mu)^2}{2\sigma^2}}dx_j$$

$$= \mu^2$$

$$\Rightarrow \langle x_i\hat{\mu}\rangle = \frac{\sigma^2}{n}+\frac{\mu^2}{n}+\frac{n-1}{n}\mu^2 = \frac{\sigma^2}{n}+\mu^2$$

$$\langle\hat{\mu}^2\rangle = \langle\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}x_ix_j\rangle$$

$$= \frac{1}{n^2}\big(n(\sigma^2+\mu^2) + (n^2-n)\mu^2\big)$$

$$= \frac{\sigma^2}{n}+\frac{\mu^2}{n}+\mu^2-\frac{\mu^2}{n} = \frac{\sigma^2}{n}+\mu^2$$

$$\Rightarrow \langle\widehat{\sigma^2}\rangle = \frac{1}{n}\sum_{i=1}^{n}\Big(\sigma^2+\mu^2-2\frac{\sigma^2}{n}-2\mu^2+\frac{\sigma^2}{n}+\mu^2\Big) = \frac{1}{n}\sum_{i=1}^{n}\Big(\sigma^2-\frac{\sigma^2}{n}\Big)$$

$$= \sigma^2\Big(1-\frac{1}{n}\Big) = \underline{\underline{\frac{n-1}{n}\sigma^2}}$$

2.

$$P(t) = \frac{1}{\tau} e^{-t/\tau} = \Gamma e^{-\Gamma t}$$

$$\mathcal{L}(\tau) = P(t_0 | \tau)$$

The likelihood based on a sample of size n is

$$\mathcal{L}(\tau) = \prod_{i=1}^{n} \frac{1}{\tau} e^{-t_i/\tau}$$

$$= \frac{1}{\tau^n} e^{-1/\tau \sum_{i=1}^{n} t_i} = \frac{1}{\tau^n} e^{-n/\tau \bar{t}}$$

where $\bar{t}$ is the sample mean. Now find the maximum likelihood.

$$\frac{d}{d\tau}\mathcal{L}(\tau) = \frac{n\bar{t}}{\tau^{n+2}} e^{\frac{-n\bar{t}}{\tau}} - \frac{n}{\tau^{n+1}} e^{\frac{-n\bar{t}}{\tau}} = \frac{n}{\tau^n} e^{\frac{-n\bar{t}}{\tau}} \left(\frac{\bar{t}}{\tau^2} - \frac{1}{\tau} \right)$$

$$\mathcal{L}'(\hat{\tau}) = 0 \Rightarrow \frac{\bar{t}}{\hat{\tau}} - 1 = 0 \Rightarrow \boxed{\hat{\tau} = \bar{t}}$$

Now repeat procedure to find $\hat{\Gamma}$.

$$\mathcal{L}(\Gamma) = \prod_{i=1}^{n} \Gamma e^{-t_i \Gamma} = \Gamma^n e^{-\Gamma n \bar{t}}$$

$$\frac{d}{d\Gamma}\mathcal{L}(\Gamma) = -\Gamma^n n\bar{t} e^{-\Gamma n \bar{t}} + n\Gamma^{n-1} e^{-\Gamma n \bar{t}} = n\Gamma^n e^{-\Gamma n \bar{t}} \left(\frac{1}{\Gamma} - \bar{t} \right)$$

$$\mathcal{L}'(\hat{\Gamma}) = 0 \Rightarrow \frac{1}{\hat{\Gamma}} - \bar{t} = 0 \Rightarrow \boxed{\hat{\Gamma} = \frac{1}{\bar{t}}}$$

2.

The next question is whether or not these estimators are biased.

$$\langle \hat{\tau} \rangle = \langle \bar{t} \rangle = \frac{1}{n}\sum_{i=1}^{n} \langle t_i \rangle = \frac{1}{n}\sum_{i=1}^{n}\int_0^\infty \frac{t_i}{\tau} e^{-t_i/\tau} dt_i$$

$$j_i \equiv -t_i/\tau \Rightarrow t_i = -j_i\tau$$
$$dt_i = -\tau dj_i$$

$$\Rightarrow \langle \hat{\tau} \rangle = \frac{\tau}{n}\sum_{i=1}^{n} \underbrace{\int_0^{-\infty} j_i e^{j_i} dj_i}_{1} = \tau \Rightarrow \hat{\tau} \text{ is } \underline{\text{unbiased}}$$

$$\langle \hat{\Gamma} \rangle = \langle \tfrac{1}{\bar{t}} \rangle$$

to calculate this, we need to find the pdf for $\frac{1}{\bar{t}}$, so lets invoke the random variable theorem.

$$\xi \equiv \frac{1}{\bar{t}},\quad g = n\left(\sum_{i=1}^{n} t_i\right)^{-1},\quad p(t_1, t_2 \ldots t_n) = \prod_{i=1}^{n} \Gamma e^{-\Gamma t_i}$$

$$\Rightarrow p(\xi) = \int_0^\infty dt_1 \int_0^\infty dt_2 \ldots \int_0^\infty dt_n\, \delta\left(\xi - n\left(\sum_{i=1}^{n} t_i\right)^{-1}\right) \prod_{i=1}^{n} \Gamma e^{-\Gamma t_i}$$

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega x} d\omega \Rightarrow \delta\left(\xi - n\left(\sum_{i=1}^{n} t_i\right)^{-1}\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega\xi} e^{\frac{-n}{\sum t_i}} d\omega$$

$$\Rightarrow p(\xi) = \frac{\Gamma^n}{2\pi}\int_{-\infty}^{\infty} d\omega\, e^{i\omega\xi} \int dt_1 \ldots \int dt_n\, e^{\frac{-n}{\sum t_i}} e^{-\Gamma \sum t_i}$$

- Not finished, probably wrong strategy here!

3.

$$P(t) = \frac{1}{\tau} e^{-t/\tau}$$

calculate the confidence belt

$$\int_{t_0-\Delta t}^{t_0+\Delta t} \frac{1}{\tau} e^{-t/\tau} dt = 0.68$$

$$\Rightarrow -e^{-t/\tau} \Big|_{t_0-\Delta t}^{t_0+\Delta t} = 0.68$$

$$\Rightarrow e^{\frac{-(t_0-\Delta t)}{\tau}} - e^{\frac{-(t_0+\Delta t)}{\tau}} = 0.68$$

$$\Rightarrow \ln\left(\frac{e^{\frac{-(t_0-\Delta t)}{\tau}}}{e^{\frac{-(t_0+\Delta t)}{\tau}}}\right) = \ln(0.68)$$

$$\Rightarrow \ln\left(e^{\frac{2\Delta t}{\tau}}\right) = \ln(0.68)$$

$$\Rightarrow \Delta t = \frac{\tau}{2}\ln(0.68)$$

- so the upper half of the belt is

$$t_0 = \tau - \frac{\tau}{2}\ln(0.68) \Rightarrow \tau_{up} = \frac{1}{1-\frac{1}{2}\ln(0.68)} t_0$$

- and the lower edge is

$$\tau_{dn} = \frac{1}{1+\frac{1}{2}\ln(0.68)} t_0$$

if $t = 4s$, the estimate is

$$\tau = 4^{+0.95}_{-0.65} \text{ s}$$

3.
Now for the 90% confidence upper limit.

$$\int_0^{\tau_u} \frac{1}{\tau} e^{-t/\tau} dt = 0.9$$

$$\Rightarrow -e^{-t/\tau}\Big|_0^{\tau_u} = 0.9$$

$$\Rightarrow 1 - e^{-\tau_u/\tau} = 0.9 \Rightarrow \ln(0.1) = -\frac{\tau_u}{\tau} \Rightarrow \tau_u = -\tau \ln(0.1)$$

$$\tau_u = \tau \ln(10)$$

so for $t_0 = 4$, $\tau_u = 9.21$

p4.

```python
#!/usr/bin/env python3
""" LPC stats HW2, Problem 4
    Author: Caleb Fangmeier
    Created: Oct. 8, 2017
"""
from scipy.stats import norm, uniform, cauchy, mode
from numpy import logspace, zeros, max, min, mean, median, var
import matplotlib.pyplot as plt


distributions = {
    'uniform': lambda a, N: uniform.rvs(loc=a-0.5, scale=1, size=N),
    'gaussian': lambda a, N: norm.rvs(loc=a, scale=1, size=N),
    'cauchy': lambda a, N: cauchy.rvs(loc=a, size=N)
}

estimators = {
    'midrange': lambda xs: max(xs) - min(xs),
    'mean': lambda xs: mean(xs),
    'median': lambda xs: median(xs),
    'mode': lambda xs: mode(xs).mode[0]
}


def var_of_est(dist_name, est_name, N, a=1):
    M = 500
    estimates = zeros(M)
    for i in range(M):  # run M experiments to estimate variance
        data = distributions[dist_name](a, N)
        estimates[i] = estimators[est_name](data)
    return var(estimates)


plt.figure()
for i, distribution in enumerate(distributions):
    plt.subplot(2, 2, i+1)
    for estimator in estimators:
        Ns = logspace(1, 3, 30, dtype=int)
        vars = zeros(30)
        for i, N in enumerate(Ns):
            vars[i] = var_of_est(distribution, estimator, N)
        plt.plot(Ns, vars, label=estimator)
    plt.title(distribution)
    plt.xlabel('N')
    plt.ylabel(r'$\sigma^2$')
    plt.xscale('log')
    plt.yscale('log')
    plt.tight_layout()
    plt.legend()

plt.show()
```

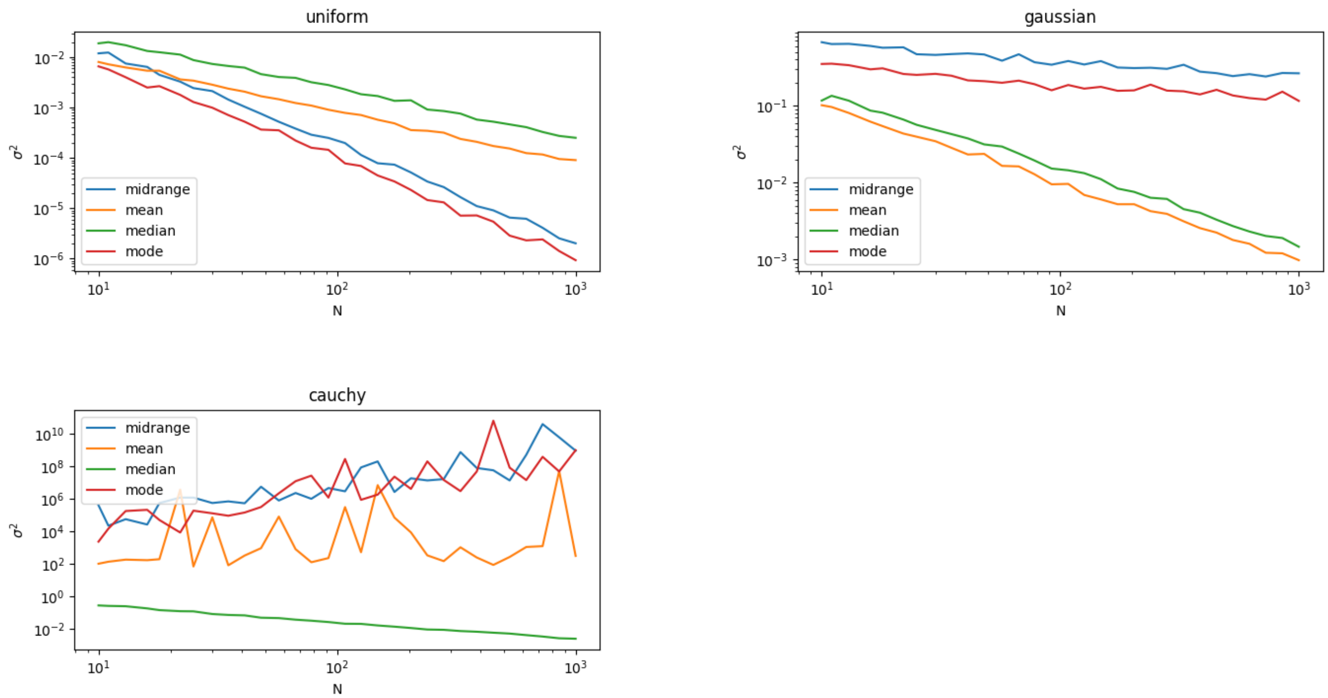

Figure 1: Variance of selected estimators for Uniform, Gaussian, and Cauchy distributions

The best estimator is different for each distribution. For the uniform distribution, the midrange and mode estimators are both good, with the mode being somewhat better. However, for the Gaussian, these estimators are beat out substantially by the median and mean estimators, with the mean being the best. Finally, for the Cauchy distribution, the median estimator is clearly the best, while all of the others fall prey to the pathological nature of Cauchy and the relatively high probability of getting sample values from far out in the tails.