11: EXAM 1 REVIEW

Stat250 S25

Prof Amanda Luby

1 General Topics

Terminology
- Parameter vs Estimator
- Estimator vs Estimate
- pdf vs likelihood
- sample vs population

Sampling distributions
- Conceptual understanding of what they are
- Central Limit Theorem for sample means
- Relationship to estimator

Estimation
- Finding the MLE
- Finding the MoM

Evaluating estimators
- Given a pdf, find the expected value and variance
- Recognize named distributions and use expected value and variance from cheat sheet
- Use properties of expected value and variance
- Comment on bias and efficiency of an estimator:
  - using the theoretical expected value and variance
  - using results from a simulation
- Determine the Cramer-Rao lower bound for a probability distribution
- Compare two estimators' efficiency in terms of sample size
- Determine if an estimator is consistent
- Use large-sample properties of the MLE

Math techniques
- Integration: polynomials, e^x, $\ln x$, and less-involved composite functions (ie $\int \ln 2x$)
- Derivatives: same as integration

R techniques
- Given simulation code and output, draw conclusions
- Given a graph and a description, draw conclusions

2 Example problems

Note: this is not an exhaustive problem list, nor is it representative of the length of the exam. Make sure to review daily prep questions, in-class examples and exercises, and homework problems.

Let $X_1, ..., X_n$ be an iid sample from a distribution with pdf $f(x|\theta) = \theta x^{\theta-1}$ for $0 < x < 1$ and $\theta > 0$.

1. Show that the likelihood function is $\theta^n \prod X_i^{\theta-1}$
2. Show that the maximum likelihood estimator is $\hat{\theta} = \frac{-n}{\sum \ln X_i}$
3. It can be shown that $E[\hat{\theta}_{MLE}] = \frac{n}{n-1}\theta$. Use this information to construct an unbiased estimator for θ, $\hat{\theta}_2$, that is based on $\hat{\theta}_{MLE}$
4. Find the method of moments estimator for θ
5. Find $I(\theta)$
6. Find the Cramer-Rao Lower Bound
7. What does the Cramer-Rao Lower Bound in (6) tell you about $Var(\hat{\theta}_{MLE})$? What about $Var(\hat{\theta}_2)$?
8. Below is a simulation comparing the performance of $\hat{\theta_{MLE}}$ to $\hat{\theta_{MoM}}$. Which output corresponds to the bias of each estimator? Which corresponds to the efficiency?

```
theta <- 2
n <- 20
n_sims <- 10000
theta_mle <- numeric(n_sims)
theta_mom <- numeric(n_sims)

for(i in 1:n_sims){
  sample <- rbeta(n, theta, 1)
  theta_mle[i] <- -n/sum(log(sample))
  theta_mom[i] <- mean(sample)/(1-mean(sample))
}

mean(theta_mle)
```

```
[1] 2.094769
```

```
var(theta_mle)
```

```
[1] 0.2464843
```

```
mean(theta_mom)
```

```
[1] 2.069172
```

```
var(theta_mom)
```

```
[1] 0.2686779
```

Are the following statements true? If not, can you correct them?

1. The variance of an estimator quantifies how close the estimator is expected to be to the true parameter value on average.
2. The Cramér-Rao Lower Bound provides an upper limit on the variance that any unbiased estimator can achieve.
3. For a given statistical model, the Method of Moments estimator is generally more statistically efficient than the Maximum Likelihood Estimator, especially for large sample sizes.
4. If an estimator is biased, it cannot be a consistent estimator.
5. Efficiency is primarily concerned with minimizing the bias of an estimator.
6. A biased estimator can sometimes have a smaller Mean Squared Error (MSE) than an unbiased estimator.

Use the graphs below for the next few questions:

(A) Sketch a sampling distribution for an estimator that is **unbiased**
(B) Sketch a sampling distribution for an estimator that is **more efficient** than the one that I sketched
(C) Sketch a sampling distribution for an estimator with **lower MSE** than the one that I sketched