Lecture Summary

Recap: Unbiased Estimators

- Let $X_1, \ldots, X_n$ be a random sample from $Expo(\lambda)$.
- Find $\hat{\lambda}_{MLE}$.
- Find $\hat{\lambda}_{MOM}$.
- Are they unbiased?
- Find the MSE of each estimator
- Are they consistent?

$\hat{\lambda}_{MLE}$

- $\hat{\lambda}_{MLE}$ is a function of $X_1, \ldots, X_n$.
- It has a sampling distribution that depends on the value of λ.
- Assume that you want to estimate how probable it is that you will make a mistake of more than 0.1.
- You want to estimate $P(|\hat{\lambda}_{MLE} - \lambda|\lambda)$
- For every possible λ, you can compute $P(|\hat{\lambda}_{MLE} - \lambda|\lambda)$ based on the sampling distribution of $\sum_{i=1}^{n} X_i$

Sampling Distributions

- Suppose $\mathrm{X} = (X_1, \ldots, X_n)$ is a random sample from $f(x|\theta)$.
- A function $r(X_1, \ldots, X_n)$ is a statistic.
- A sampling distribution: the distribution of a statistic (given θ).
- Can use the sampling distributions to compare different estimators and to determine the sample size we need.
- Used to get confidence intervals and to do hypothesis testing.
- Leads to definitions of new distributions, e.g. χ^2_m and t_m.

The χ^2 distributions

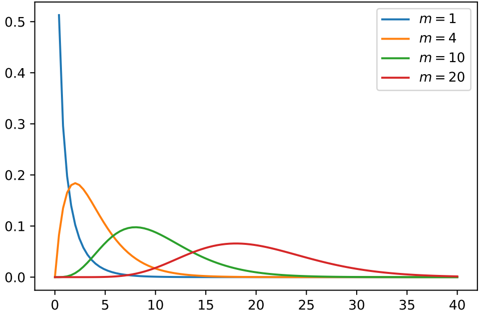
Definition

The χ^2 distribution with m degrees of freedom is the $Gamma(a = m/2, \beta = 1/2)$. The pdf is

$$f(x|m) = \frac{1}{2^{m/2}\Gamma(m/2)} x^{m/2-1} e^{-x/2}$$

If $X \sim \chi^2_m$ then

- $E(X) = m$
- $Var(X) = 2m$



Properties of the χ^2 distributions

Theorem (8.2.1)

Let $X_1, \ldots, X_n$ be independent random variables and $X_i \sim \chi^2_{m_i}$. Then

$$X_1 + \cdots + X_n \sim \chi^2_m$$

where $m = m_1 + \cdots + m_n$.

- follows from Theorem 5.7.7.

Theorem (8.2.2)

If $X \sim \mathcal{N}(0, 1)$, then $X^2 \sim \chi^2_1$

Properties of the χ^2 distributions

Corollary (8.2.1)

If the random variables $X_1, \ldots, X_n$ i.i.d., $X_i \sim \mathcal{N}(0,1)$ then

$$X_1^2 + \cdots + X_n^2 \sim \chi_n^2$$

The χ_m^2 distribution is a sampling distribution related to the sample variance of a normal distribution:

- If $X_1, \ldots, X_n$ are i.i.d, $X_i \sim \mathcal{N}(\mu, \sigma^2)$, where μ is known and the MLE of σ^2 is

$$\hat{\sigma^2_{MLE}} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

- Then

$$\frac{n\hat{\sigma^2_{MLE}}}{\sigma^2} \sim \chi_n^2$$

Sampling Distributions of the normal sample mean and variance

- Let $X_1, \ldots, X_n$ be a random sample from a $\mathcal{N}(\mu, \sigma^2)$ with unknown μ, σ^2.
- The sample mean and the sample variance are defined as

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad S_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

- They are the MLEs for μ and σ^2 in this setting.

Theorem (8.3.1)

Let $X_1, \ldots, X_n$ be a random sample from $\mathcal{N}(\mu, \sigma^2)$. Then $\bar{X}_n$ and S_n are independent random variables and $\bar{X}_n \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$, $\frac{nS_n}{\sigma^2} \sim \chi^2_{n-1}$.

Sampling Distribution of the sample mean and variance

Theorem (8.3.1)

Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Then $\overline{X}_n$ and S_n are independent random variables and

$$\overline{X_n} \sim \mathcal{N}(\mu, \sigma^2/n),$$

$$\frac{n}{\sigma^2} S_n = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2 \sim \chi^2_{n-1}$$

- Replacing μ by $\overline{X}_n$ results in one less degree of freedom.
- $\overline{X}_n$ and S_n are functions of the same random variables, but they are independent (this only happens when the random sample is drawn from a normal distribution).

Recap

- Today we discussed samples of normal distributions:
- MLE estimators of μ and σ^2 in different settings.
- Sampling distributions of these estimators.
- Sample mean and sample variance are independent! (see an example in section 8.3 of your book)

Question: Bounding errors in estimates

- Suppose that $X_1, \ldots, X_m$ form a random sample from the normal distribution with unknown mean μ and unknown variance σ^2.
- Assuming that the sample size n is 16, determine the values of the following probabilities:

$$P(0.5\sigma^2 \leq \hat{\sigma}_0^2 \leq 2\sigma^2)$$

$$P(0.5\sigma^2 \leq S_n \leq 2\sigma^2)$$

- Can you find the smallest valueof n such that

$$P(0.5\sigma^2 \leq S_n \leq 2\sigma^2) \geq 0.9?$$

Practice Exercises

8.1	1, 2
8.2	1, 4, 9
8.3	1, 6, 8