The method of moments

Chrysafis Vogiatzis

Department of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana-Champaign

Lecture 17

Last time..

Last time, we saw what makes a good point estimator $\hat{\Theta}$:

- small **bias** (zero preferably). $bias = E\left[\hat{\Theta}\right] - \theta.$
- small **variance** (minimum among all estimators). $Var\left[\hat{\Theta}\right].$
- small **mean square error**. $MSE = bias^2 + Var\left[\hat{\Theta}\right].$
- We can also define the **relative efficiency** of $\hat{\Theta}_1, \hat{\Theta}_2$: $\frac{MSE\left(\hat{\Theta}_1\right)}{MSE\left(\hat{\Theta}_2\right)}$

Today, we begin a discussion on methods to obtain good estimators with:

- The method of moments.

Last time..

Last time, we saw what makes a good point estimator $\hat{\Theta}$:

- small **bias** (zero preferably). $bias = E\left[\hat{\Theta}\right] - \theta.$
- small **variance** (minimum among all estimators). $Var\left[\hat{\Theta}\right].$
- small **mean square error.** $MSE = bias^2 + Var\left[\hat{\Theta}\right].$
- We can also define the **relative efficiency** of $\hat{\Theta}_1, \hat{\Theta}_2$: $\frac{MSE\left(\hat{\Theta}_1\right)}{MSE\left(\hat{\Theta}_2\right)}$

Today, we begin a discussion on methods to obtain good estimators with:

- The method of moments.

ILLINOIS

Last time..

Last time, we saw what makes a good point estimator $\hat{\Theta}$:

- small **bias** (zero preferably). $bias = E\left[\hat{\Theta}\right] - \theta.$
- small **variance** (minimum among all estimators). $Var\left[\hat{\Theta}\right].$
- small **mean square error**. $MSE = bias^2 + Var\left[\hat{\Theta}\right].$
- We can also define the **relative efficiency** of $\hat{\Theta}_1, \hat{\Theta}_2$: $\frac{MSE\left(\hat{\Theta}_1\right)}{MSE\left(\hat{\Theta}_2\right)}$

Today, we begin a discussion on methods to obtain good estimators with:

- The method of moments.

Last time..

Last time, we saw what makes a good point estimator $\hat{\Theta}$:

- small **bias** (zero preferably). $bias = E\left[\hat{\Theta}\right] - \theta.$
- small **variance** (minimum among all estimators). $Var\left[\hat{\Theta}\right].$
- small **mean square error**. $MSE = bias^2 + Var\left[\hat{\Theta}\right].$
- We can also define the **relative efficiency** of $\hat{\Theta}_1, \hat{\Theta}_2$: $\frac{MSE\left(\hat{\Theta}_1\right)}{MSE\left(\hat{\Theta}_2\right)}$

Today, we begin a discussion on methods to obtain good estimators with:

- The method of moments.

ILLINOIS

Last time..

Last time, we saw what makes a good point estimator $\hat{\Theta}$:

- small **bias** (zero preferably). $bias = E\left[\hat{\Theta}\right] - \theta.$
- small **variance** (minimum among all estimators). $Var\left[\hat{\Theta}\right].$
- small **mean square error**. $MSE = bias^2 + Var\left[\hat{\Theta}\right].$
- We can also define the **relative efficiency** of $\hat{\Theta}_1, \hat{\Theta}_2$: $\frac{MSE\left(\hat{\Theta}_1\right)}{MSE\left(\hat{\Theta}_2\right)}$

Today, we begin a discussion on methods to obtain good estimators with:

- The method of moments.

ILLINOIS

Moments

First things first: what do we know? Assume we have:

- Some population distributed with pdf $f(x)$.
- $f(x)$ depends on m parameters, $\theta_1, \theta_2, \ldots, \theta_m$.
- Let $X_1, X_2, \ldots, X_n$ be a sample of that population.

Then:

Definition (Population moments)

We define the k-th moment of a population X with pdf $f(x)$ as $E\left[X^k\right]$.

Definition (Sample moments)

We define the k-th moment of a sample $X_1, X_2, \ldots, X_n$ as $\frac{1}{n}\sum_{i=1}^{n} X^k$.

Moments

First things first: what do we know? Assume we have:

- Some population distributed with pdf $f(x)$.
- $f(x)$ depends on m parameters, $\theta_1, \theta_2, \ldots, \theta_m$.
- Let $X_1, X_2, \ldots, X_n$ be a sample of that population.

Then:

Definition (Population moments)

We define the k-th moment of a population X with pdf $f(x)$ as $E\left[X^k\right]$.

Definition (Sample moments)

We define the k-th moment of a sample $X_1, X_2, \ldots, X_n$ as $\frac{1}{n}\sum_{i=1}^{n} X^k$.

ILLINOIS

Moments

First things first: what do we know? Assume we have:

- Some population distributed with pdf $f(x)$.
- $f(x)$ depends on m parameters, $\theta_1, \theta_2, \ldots, \theta_m$.
- Let $X_1, X_2, \ldots, X_n$ be a sample of that population.

Then:

Definition (Population moments)

We define the k-th moment of a population X with pdf $f(x)$ as $E\left[X^k\right]$.

Definition (Sample moments)

We define the k-th moment of a sample $X_1, X_2, \ldots, X_n$ as $\frac{1}{n}\sum_{i=1}^{n} X^k$.

ILLINOIS

Moments

First things first: what do we know? Assume we have:

- Some population distributed with pdf $f(x)$.
- $f(x)$ depends on m parameters, $\theta_1, \theta_2, \ldots, \theta_m$.
- Let $X_1, X_2, \ldots, X_n$ be a sample of that population.

Then:

Definition (Population moments)

We define the k-th moment of a population X with pdf $f(x)$ as $E\left[X^k\right]$.

Definition (Sample moments)

We define the k-th moment of a sample $X_1, X_2, \ldots, X_n$ as $\frac{1}{n}\sum_{i=1}^{n} X^k$.

ILLINOIS

Moments

First things first: what do we know? Assume we have:

- Some population distributed with pdf $f(x)$.
- $f(x)$ depends on m parameters, $\theta_1, \theta_2, \ldots, \theta_m$.
- Let $X_1, X_2, \ldots, X_n$ be a sample of that population.

Then:

Definition (Population moments)

We define the k-th moment of a population X with pdf $f(x)$ as $E\left[X^k\right]$.

Definition (Sample moments)

We define the k-th moment of a sample $X_1, X_2, \ldots, X_n$ as $\frac{1}{n}\sum_{i=1}^{n} X^k$.

ILLINOIS

Method of moments

A few notes:

- The k-th moment of $f(x)$ (calculate as $E\left[X^k\right]$) depends only on the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$.
- The k-th moment of the sample, $\frac{1}{n}\sum_{i=1}^{n} X_i^k$ depends only on the data and can be assigned a numeric value.

Moment estimators $\hat{\Theta}_1, \hat{\Theta}_2, \ldots, \hat{\Theta}_m$ for each of the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$ can be obtained following the procedure:

1. Get the first m moments of $f(x)$ and of the sample.
2. Equate them.
3. Solve a system of equations with m unknowns (the parameters θ_i) to obtain the **moment estimators**.

ILLINOIS

Method of moments

A few notes:

- The k-th moment of $f(x)$ (calculate as $E\left[X^k\right]$) depends only on the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$.
- The k-th moment of the sample, $\frac{1}{n}\sum_{i=1}^{n} X_i^k$ depends only on the data and can be assigned a numeric value.

Moment estimators $\hat{\Theta}_1, \hat{\Theta}_2, \ldots, \hat{\Theta}_m$ for each of the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$ can be obtained following the procedure:

1. Get the first m moments of $f(x)$ and of the sample.
2. Equate them.
3. Solve a system of equations with m unknowns (the parameters θ_i) to obtain the **moment estimators**.

ILLINOIS

Method of moments

A few notes:

- The k-th moment of $f(x)$ (calculate as $E\left[X^k\right]$) depends only on the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$.
- The k-th moment of the sample, $\frac{1}{n}\sum_{i=1}^{n} X_i^k$ depends only on the data and can be assigned a numeric value.

Moment estimators $\hat{\Theta}_1, \hat{\Theta}_2, \ldots, \hat{\Theta}_m$ for each of the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$ can be obtained following the procedure:

1. Get the first m moments of $f(x)$ and of the sample.
2. Equate them.
3. Solve a system of equations with m unknowns (the parameters θ_i) to obtain the **moment estimators**.

ILLINOIS

Method of moments

A few notes:

- The k-th moment of $f(x)$ (calculate as $E\left[X^k\right]$) depends only on the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$.
- The k-th moment of the sample, $\frac{1}{n}\sum_{i=1}^{n} X_i^k$ depends only on the data and can be assigned a numeric value.

Moment estimators $\hat{\Theta}_1, \hat{\Theta}_2, \ldots, \hat{\Theta}_m$ for each of the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$ can be obtained following the procedure:

1. Get the first m moments of $f(x)$ and of the sample.
2. Equate them.
3. Solve a system of equations with m unknowns (the parameters θ_i) to obtain the **moment estimators**.

ILLINOIS

Method of moments

A few notes:

- The k-th moment of $f(x)$ (calculate as $E\left[X^k\right]$) depends only on the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$.
- The k-th moment of the sample, $\frac{1}{n}\sum_{i=1}^{n} X_i^k$ depends only on the data and can be assigned a numeric value.

Moment estimators $\hat{\Theta}_1, \hat{\Theta}_2, \ldots, \hat{\Theta}_m$ for each of the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$ can be obtained following the procedure:

1. Get the first m moments of $f(x)$ and of the sample.
2. Equate them.
3. Solve a system of equations with m unknowns (the parameters θ_i) to obtain the **moment estimators**.

ILLINOIS

Method of moments

A few notes:

- The k-th moment of $f(x)$ (calculate as $E\left[X^k\right]$) depends only on the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$.
- The k-th moment of the sample, $\frac{1}{n}\sum_{i=1}^{n} X_i^k$ depends only on the data and can be assigned a numeric value.

Moment estimators $\hat{\Theta}_1, \hat{\Theta}_2, \ldots, \hat{\Theta}_m$ for each of the unknown parameters $\theta_1, \theta_2, \ldots, \theta_m$ can be obtained following the procedure:

1. Get the first m moments of $f(x)$ and of the sample.
2. Equate them.
3. Solve a system of equations with m unknowns (the parameters θ_i) to obtain the **moment estimators**.

ILLINOIS

Method of moments: example

Example

Suppose we have been observing the times between accidents in a factory that we suspect are exponentially distributed. We have collected the following times so far: $X_1 = 3$ *days,* $X_2 = 4$ *days,* $X_3 = 2$ *days,* $X_4 = 3$ *days,* $X_5 = 2$ *days. What is the rate* λ*?*

Answer: One unknown, so only one moment needed:

- Population 1st moment:
- Sample 1st moment:

We have:

$$\frac{1}{\lambda} = 2.8 \text{ days} \implies \hat{\lambda} = 1 \text{ accident every } 2.8 \text{ days.}$$

We write it as $\hat{\lambda}$ in the end since it is an estimator for the *true* value of λ.

Method of moments: example

Example

Suppose we have been observing the times between accidents in a factory that we suspect are exponentially distributed. We have collected the following times so far: $X_1 = 3$ *days,* $X_2 = 4$ *days,* $X_3 = 2$ *days,* $X_4 = 3$ *days,* $X_5 = 2$ *days. What is the rate* λ*?*

Answer: One unknown, so only one moment needed:

- Population 1st moment: $E\left[X^1\right] = E\left[X\right] = \frac{1}{\lambda}$
- Sample 1st moment: $\frac{1}{5}\sum_{i=1}^{5} X_i^1 = \frac{1}{5}(3+4+2+3+2) = 2.8$ days.

We have:

$$\frac{1}{\lambda} = 2.8 \text{ days} \implies \hat{\lambda} = 1 \text{ accident every } 2.8 \text{ days.}$$

We write it as $\hat{\lambda}$ in the end since it is an estimator for the *true* value of λ.

Method of moments: example

Example

Suppose we have been observing the times between accidents in a factory that we suspect are exponentially distributed. We have collected the following times so far: $X_1 = 3$ *days,* $X_2 = 4$ *days,* $X_3 = 2$ *days,* $X_4 = 3$ *days,* $X_5 = 2$ *days. What is the rate* λ*?*

Answer: One unknown, so only one moment needed:

- Population 1st moment: $E\left[X^1\right] = E[X] = \frac{1}{\lambda}$
- Sample 1st moment: $\frac{1}{5}\sum_{i=1}^{5} X_i^1 = \frac{1}{5}(3+4+2+3+2) = 2.8$ days.

We have:

$$\frac{1}{\lambda} = 2.8 \text{ days} \implies \hat{\lambda} = 1 \text{ accident every 2.8 days.}$$

We write it as $\hat{\lambda}$ in the end since it is an estimator for the *true* value of λ.

Method of moments: example

Example

Suppose we have been observing the times between accidents in a factory that we suspect are exponentially distributed. We have collected the following times so far: $X_1 = 3$ *days,* $X_2 = 4$ *days,* $X_3 = 2$ *days,* $X_4 = 3$ *days,* $X_5 = 2$ *days. What is the rate* λ*?*

Answer: One unknown, so only one moment needed:

- Population 1st moment: $E\left[X^1\right] = E\left[X\right] = \frac{1}{\lambda}$
- Sample 1st moment: $\frac{1}{5}\sum_{i=1}^{5} X_i^1 = \frac{1}{5}\left(3+4+2+3+2\right) = 2.8$ days.

We have:

$$\frac{1}{\lambda} = 2.8 \text{ days} \implies \hat{\lambda} = 1 \text{ accident every 2.8 days.}$$

We write it as $\hat{\lambda}$ in the end since it is an estimator for the *true* value of λ.

ILLINOIS

Method of moments: example

Example

Suppose we have been observing the times between accidents in a factory that we suspect are exponentially distributed. We have collected the following times so far: $X_1 = 3$ days, $X_2 = 4$ days, $X_3 = 2$ days, $X_4 = 3$ days, $X_5 = 2$ days. What is the rate λ?

Answer: One unknown, so only one moment needed:

- Population 1st moment: $E\left[X^1\right] = E[X] = \frac{1}{\lambda}$
- Sample 1st moment: $\frac{1}{5}\sum_{i=1}^{5} X_i^1 = \frac{1}{5}(3+4+2+3+2) = 2.8$ days.

We have:

$$\frac{1}{\lambda} = 2.8 \text{ days} \implies \hat{\lambda} = 1 \text{ accident every } 2.8 \text{ days.}$$

We write it as $\hat{\lambda}$ in the end since it is an estimator for the *true* value of λ.

ILLINOIS

Method of moments: example

Example

Suppose we have been observing the times between accidents in a factory that we suspect are exponentially distributed. We have collected the following times so far: $X_1 = 3$ days, $X_2 = 4$ days, $X_3 = 2$ days, $X_4 = 3$ days, $X_5 = 2$ days. What is the rate λ?

Answer: One unknown, so only one moment needed:

- Population 1st moment: $E\left[X^1\right] = E\left[X\right] = \frac{1}{\lambda}$
- Sample 1st moment: $\frac{1}{5}\sum_{i=1}^{5} X_i^1 = \frac{1}{5}(3+4+2+3+2) = 2.8$ days.

We have:

$$\frac{1}{\lambda} = 2.8 \text{ days} \implies \hat{\lambda} = 1 \text{ accident every } 2.8 \text{ days.}$$

We write it as $\hat{\lambda}$ in the end since it is an estimator for the *true* value of λ.

ILLINOIS

Method of moments: another example

Example

We believe the times it takes to deliver a package are normally distributed with unknown μ and σ^2. We have collected information on 10 packages and the time to delivery (in hours) are: 49.1, 47.9, 48.6, 50.4, 49.5, 49.8, 48.2, 50.3, 45.2, 46.2. What are good mean and variance estimators for the normal distribution using the method of moments?

Answer: Two unknown (mean and variance), so we need two moments:

- Population 1st moment:
- Sample 1st moment:
- Population 2nd moment:
- Sample 2nd moment:

Equating, we get $\hat{\mu} = 48.52$ and $\hat{\sigma}^2 = 2.6536$.

ILLINOIS

Method of moments: another example

Example

We believe the times it takes to deliver a package are normally distributed with unknown μ and σ^2. We have collected information on 10 packages and the time to delivery (in hours) are: 49.1, 47.9, 48.6, 50.4, 49.5, 49.8, 48.2, 50.3, 45.2, 46.2. What are good mean and variance estimators for the normal distribution using the method of moments?

Answer: Two unknown (mean and variance), so we need two moments:

- Population 1st moment:

$$E\left[X^1\right] = E[X] = \mu$$

- Sample 1st moment:

$$\frac{1}{10}\sum_{i=1}^{10} X_i^1 = 48.52$$

- Population 2nd moment:

$$E\left[X^2\right] = Var[X] - (E[X])^2 = \sigma^2 - \mu^2$$

- Sample 2nd moment:

$$\frac{1}{10}\sum_{i=1}^{10} X_i^2 = 2356.844$$

Equating, we get $\hat{\mu} = 48.52$ and $\hat{\sigma}^2 = 2.6536$.

ILLINOIS

Method of moments: another example

Example

We believe the times it takes to deliver a package are normally distributed with unknown μ and σ^2. We have collected information on 10 packages and the time to delivery (in hours) are: 49.1, 47.9, 48.6, 50.4, 49.5, 49.8, 48.2, 50.3, 45.2, 46.2. What are good mean and variance estimators for the normal distribution using the method of moments?

Answer: Two unknown (mean and variance), so we need two moments:

- Population 1st moment:

$$E\left[X^1\right] = E[X] = \mu$$

- Sample 1st moment:

$$\frac{1}{10}\sum_{i=1}^{10} X_i^1 = 48.52$$

- Population 2nd moment:

$$E\left[X^2\right] = Var[X] - (E[X])^2 = \sigma^2 - \mu^2$$

- Sample 2nd moment:

$$\frac{1}{10}\sum_{i=1}^{10} X_i^2 = 2356.844$$

Equating, we get $\hat{\mu} = 48.52$ and $\hat{\sigma}^2 = 2.6536$.

Method of moments: another example

Example

We believe the times it takes to deliver a package are normally distributed with unknown μ and σ^2. We have collected information on 10 packages and the time to delivery (in hours) are: 49.1, 47.9, 48.6, 50.4, 49.5, 49.8, 48.2, 50.3, 45.2, 46.2. What are good mean and variance estimators for the normal distribution using the method of moments?

Answer: Two unknown (mean and variance), so we need two moments:

- Population 1st moment:

$$E\left[X^1\right] = E[X] = \mu$$

- Sample 1st moment:

$$\frac{1}{10}\sum_{i=1}^{10} X_i^1 = 48.52$$

- Population 2nd moment:

$$E\left[X^2\right] = Var[X] - (E[X])^2 = \sigma^2 - \mu^2$$

- Sample 2nd moment:

$$\frac{1}{10}\sum_{i=1}^{10} X_i^2 = 2356.844$$

Equating, we get $\hat{\mu} = 48.52$ and $\hat{\sigma}^2 = 2.6536$.

ILLINOIS

Method of moments: another example

Example

We believe the times it takes to deliver a package are normally distributed with unknown μ and σ^2. We have collected information on 10 packages and the time to delivery (in hours) are: 49.1, 47.9, 48.6, 50.4, 49.5, 49.8, 48.2, 50.3, 45.2, 46.2. What are good mean and variance estimators for the normal distribution using the method of moments?

Answer: Two unknown (mean and variance), so we need two moments:

- Population 1st moment:

$$E\left[X^1\right] = E[X] = \mu$$

- Sample 1st moment:

$$\frac{1}{10}\sum_{i=1}^{10} X_i^1 = 48.52$$

- Population 2nd moment:

$$E\left[X^2\right] = Var[X] - (E[X])^2 = \sigma^2 - \mu^2$$

- Sample 2nd moment:

$$\frac{1}{10}\sum_{i=1}^{10} X_i^2 = 2356.844$$

Equating, we get $\hat{\mu} = 48.52$ and $\hat{\sigma}^2 = 2.6536$.

ILLINOIS

Method of moments: recap

- Given a population X with pdf $f(x)$ and $\theta_1, \theta_2, \ldots, \theta_m$ are some unknown parameters.
- Define population moments as $E\left[X^k\right]$ and sample moments as $\frac{1}{n}\sum_{i=1}^{n} X_i^k$, for $k \geq 1$.
- Take the first m moments and equate them.
- The system solution gives us the so-called **moment estimators** $\hat{\Theta}_1, \hat{\Theta}_2, \ldots, \hat{\Theta}_m$ for the m unknown parameters.

ILLINOIS