

Central Limit Theorem

If X_1, X_2, \dots, X_n is a random sample of size n is taken from a population with mean μ and **finite variance σ^2** , and **any distribution**. If \bar{X} is the sample mean, then the **limiting form of the distribution** of

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad (7-1)$$

for **large n** , is the **standard normal distribution**.

If X_1, X_2, \dots, X_n are themselves normally distributed – for any n

Example 7-1: Resistors

An electronics company manufactures resistors having a mean resistance of 100 ohms and a standard deviation of 10 ohms. What is the approximate probability that a random sample of $n = 25$ resistors will have an average resistance of less than 95 ohms?

Example 7-1: Resistors

An electronics company manufactures resistors having a mean resistance of 100 ohms and a standard deviation of 10 ohms. What is the approximate probability that a random sample of $n = 25$ resistors will have an average resistance of less than 95 ohms?

$$\mu = 100 \text{ ohms}, \quad \sigma = 10 \text{ ohms}, \quad n = 25$$

$$\mu_{\bar{x}} = \mu; \quad \sigma_{\bar{x}_n} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = \frac{10}{5} = 2 \text{ ohms}$$

$$Z_{\bar{x}} = \frac{95 - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{95 - 100}{2} = -2.5$$

$$\begin{aligned} \text{Prob}(\bar{X} < 95) &= \Phi(Z_{\bar{x}}) = \Phi(-2.5) = \\ &= 0.0062 \end{aligned}$$

Example 7-1: Resistors

An electronics company manufactures resistors having a mean resistance of 100 ohms and a standard deviation of 10 ohms. What is the approximate probability that a random sample of $n = 25$ resistors will have an average resistance of less than 95 ohms?

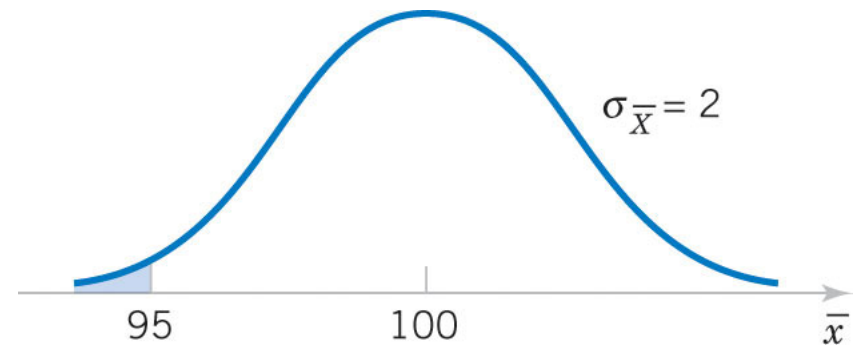


Figure 7-2 Desired probability is shaded

Answer:

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2.0$$
$$\Phi\left(\frac{\bar{X} - \mu}{\sigma_{\bar{X}}}\right) = \Phi\left(\frac{95 - 100}{2}\right)$$
$$= \Phi(-2.5) = 0.0062$$

Two Populations

We have two independent populations. What is the distribution of the difference of their sample means?

The sampling distribution of $\bar{X}_1 - \bar{X}_2$ has the following mean and variance:

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2$$

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

Sampling Distribution of a Difference in Sample Means

- **If** we have two independent populations with means μ_1 and μ_2 , and variances σ_1^2 and σ_2^2 ,
- **And if** \bar{X}_1 and \bar{X}_2 are the sample means of two independent random samples of sizes n_1 and n_2 from these populations:
- **Then** the sampling distribution of:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (7-4)$$

is approximately standard normal, if the conditions of the central limit theorem apply.

- **If** the two populations are normal, **then** the sampling distribution is exactly standard normal.

Example 7-3: Aircraft Engine Life

The effective life of a component used in jet-turbine aircraft engines is a random variable with $\mu_{\text{old}}=5000$ hours and $\sigma_{\text{old}}=40$ hours (old). The engine manufacturer introduces an improvement into the manufacturing process for this component that changes the parameters to $\mu_{\text{new}}=5050$ hours and $\sigma_{\text{new}}=30$ hours (new).

Random samples of 16 components manufactured using “old” process and 25 components using “new” process are chosen.

What is the probability new sample mean is at least 25 hours longer than old?

Example 7-3: Aircraft Engine Life

The effective life of a component used in jet-turbine aircraft engines is a random variable with $\mu_{old}=5000$ hours and $\sigma_{old}=40$ hours (old). The engine manufacturer introduces an improvement into the manufacturing process for this component that changes the parameters to $\mu_{new}=5050$ hours and $\sigma_{new}=30$ hours (new).

Random samples of 16 components manufactured using "old" process and 25 components using "new" process are chosen.

What is the probability new sample mean is at least 25 hours longer than old?

$$\sigma_{\bar{X}_{old}} = \frac{\sigma_{old}}{\sqrt{16}} = 10 \text{ hrs}$$

$$\sigma_{\bar{X}_{new}} = \frac{\sigma_{new}}{\sqrt{25}} = 6 \text{ hrs}$$

$$\sigma_{TOT} = \sqrt{\sigma_{\bar{X}_{old}}^2 + \sigma_{\bar{X}_{new}}^2} =$$

$$= \sqrt{100 + 36} \approx 11.7 \text{ hrs}$$

$$\mu_{new} - \mu_{old} = 50 \text{ hrs}$$

$$z = \frac{25 - (50)}{11.7} = -2.14$$

$$\text{Prob}(z > -2.14) = 0.9840$$

Example 7-3: Aircraft Engine Life

The effective life of a component used in jet-turbine aircraft engines is a normal-distributed random variable with parameters shown (old). The engine manufacturer introduces an improvement into the manufacturing process for this component that changes the parameters μ and σ as shown (new).

Random samples are selected from the “old” process and “new” process as shown.

What is the probability new sample mean is at least 25 hours longer than old?

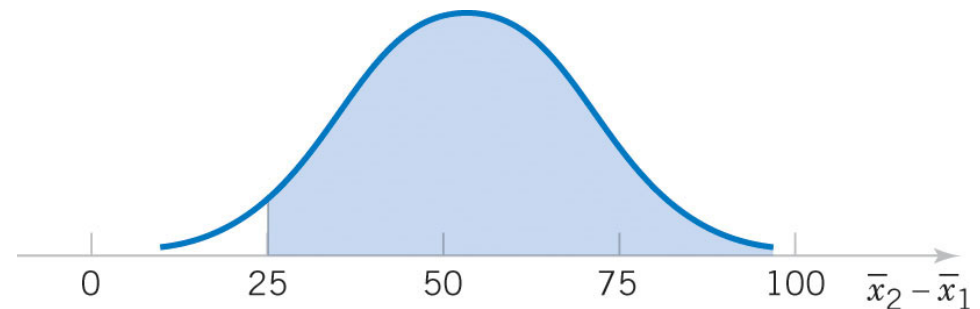


Figure 7-4 Sampling distribution of the sample mean difference.

	Process		
	Old (1)	New (2)	Diff (2-1)
$\mu =$	5,000	5,050	50
$\sigma =$	40	30	50
$n =$	16	25	
Calculations			
$s / \sqrt{n} =$	10	6	11.7
		$z =$	-2.14
$P(\bar{x}_2 - \bar{x}_1 > 25) = P(Z > z) =$			0.9840

Credit: XKCD
comics

WHY ARE THERE SLAVES IN THE BIBLE

WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS

WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT

QUESTIONS

FOUND IN GOOGLE AUTOCOMplete



WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY

WHY AREN'T ECONOMISTS RICH

WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY IS THERE PHLEGM

WHY DO AMERICANS CALL IT SOCCER

WHY IS PSYCHIC WEAK TO BUG

WHY ARE MY EARS RINGING

WHY DO CHILDREN GET CANCER

WHY ARE THERE SO MANY AVENGERS

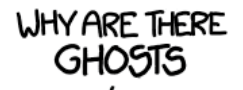
WHY IS POSEIDON ANGRY WITH ODYSSEUS

WHY ARE THE AVENGERS FIGHTING THE X MEN

WHY IS THERE ICE IN SPACE

WHY ARE THERE ANTS IN MY LAPTOP

WHY IS EARTH TILTED



WHY IS THERE AN OWL IN MY BACKYARD

WHY IS SPACE BLACK



WHY IS THERE AN OWL OUTSIDE MY WINDOW

WHY IS OUTER SPACE SO COLD

WHY IS THERE AN OWL ON THE DOLLAR BILL

WHY ARE THERE PYRAMIDS ON THE MOON

WHY DO OWLS ATTACK PEOPLE

WHY IS NASA SHUTTING DOWN

WHY ARE AK 47s SO EXPENSIVE

WHY ARE THERE MALE AND FEMALE BIKES

WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE

WHY ARE THERE TINY SPIDERS IN MY HOUSE

WHY ARE THERE GODS

WHY DO SPIDERS COME INSIDE

WHY ARE THERE TWO SPOCKS

WHY ARE THERE HUGE SPIDERS IN MY HOUSE

WHY IS MT VESUVIUS THERE

WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE

WHY DO THEY SAY T MINUS

WHY ARE THERE SPIDERS IN MY ROOM

WHY ARE THERE OBELISKS

WHY ARE THERE SO MANY SPIDERS IN MY ROOM

WHY ARE WRESTLERS ALWAYS WET

WHY DO SPIDER BITES ITCH

WHY ARE OCEANS BECOMING MORE ACIDIC

WHY IS DYING SO SCARY

WHY IS ARWEN DYING



WHY IS LIFE SO BORING
WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA

WHY DO WHALES JUMP
WHY ARE WITCHES GREEN
WHY ARE THERE MIRRORS ABOVE BEDS

WHY DO I SAY UH
WHY IS SEA SALT BETTER
WHY ARE THERE TREES IN THE MIDDLE OF FIELDS

WHY IS THERE NOT A POKEMON MMO
WHY IS THERE LAUGHING IN TV SHOWS
WHY ARE THERE DOORS ON THE FREEWAY

WHY ARE THERE SO MANY SVCHOST.EXE RUNNING
WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA
WHY ARE THERE SCARY SOUNDS IN MINECRAFT

WHY IS THERE KICKING IN MY STOMACH
WHY ARE THERE TWO SLASHES AFTER HTTP
WHY ARE THERE CELEBRITIES

WHY DO SNAKES EXIST
WHY DO OYSTERS HAVE PEARLS
WHY ARE DUCKS CALLED DUCKS

WHY DO THEY CALL IT THE CLAP
WHY ARE KYLE AND CARTMAN FRIENDS
WHY IS THERE AN ARROW ON AANG'S HEAD

WHY ARE TEXT MESSAGES BLUE
WHY ARE THERE MUSTACHES ON CLOTHES
WHY ARE THERE MUSTACHES ON CARS

WHY ARE THERE MUSTACHES EVERYWHERE
WHY ARE THERE SO MANY BIRDS IN OHIO
WHY IS THERE SO MUCH RAIN IN OHIO

WHY IS OHIO WEATHER SO WEIRD
WHY ARE THERE BRIDESMAIDS
WHY DO DYING PEOPLE REACH UP

WHY AREN'T THERE VARICOSE ARTERIES
WHY ARE OLD KUNGONS DIFFERENT

WHY ARE THERE SQUIRRELS
WHY IS PROGRAMMING SO HARD
WHY IS THERE A 0 OHM RESISTOR

WHY DO AMERICANS HATE SOCCER
WHY DO RHYMES SOUND GOOD
WHY DO TREES DIE

WHY DO IGUANAS DIE
WHY AREN'T THERE DINOSAUR GHOSTS

WHY IS THERE HELL IF GOD FORGIVES
WHY IS GPS FREE

WHY IS SEX SO IMPORTANT



WHY ARE THERE FEMALE MR NIMES

Descriptive statistics:

Point estimation:

Point Estimation

- A sample was collected: X_1, X_2, \dots, X_n
- We suspect that sample was drawn from a random variable distribution $f(x)$
- $f(x)$ has k parameters that we do not know
- Point estimates are estimates of the parameters of the $f(x)$ describing the population based on the sample
 - For exponential PDF: $f(x) = \lambda \exp(-\lambda x)$ one wants to estimate λ
 - For Bernoulli PDF: $p^x(1-p)^{1-x}$ one wants to estimate p
 - For normal PDF one wants to estimate both μ and σ
- Point estimates are uncertain: therefore, we can talk of averages and standard deviations of point estimators

Point Estimator

A **point estimate** of some parameter θ describing population random variable is a single numerical value $\hat{\theta}$ depending on all values x_1, x_2, \dots, x_n in the sample.

The sample statistic (whis a random variable $\hat{\Theta}$ defined by a function $\hat{\Theta}(X_1, X_2, \dots, X_n)$) is called the **point estimator**.

- There could be **multiple choices** for the point estimator of a parameter.
- To estimate the **mean of a population**, we could choose the:
 - **Sample mean**
 - Sample median
 - Peak of the histogram
 - $\frac{1}{2}$ of (largest + smallest) observations of the sample.
- We need to develop criteria to compare estimates using statistical properties.

Unbiased Estimators Defined

The point estimator $\hat{\Theta}$ is an **unbiased estimator**

for the parameter θ if:

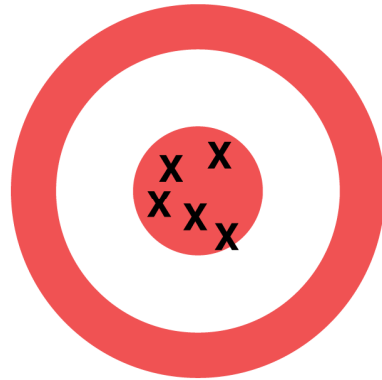
$$E(\hat{\Theta}) = \theta \quad (7-5)$$

If the estimator is not unbiased, then the difference:

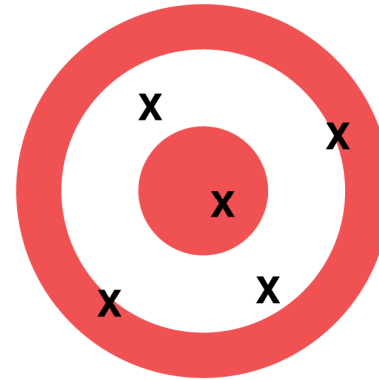
$$E(\hat{\Theta}) - \theta \quad (7-6)$$

is called the **bias** of the estimator $\hat{\Theta}$.

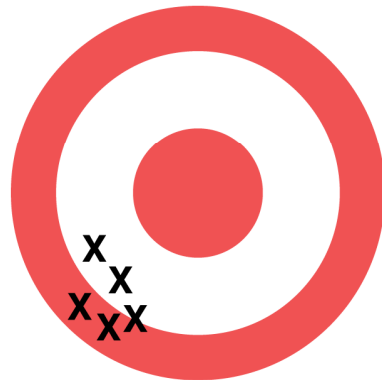
Bias vs Noise



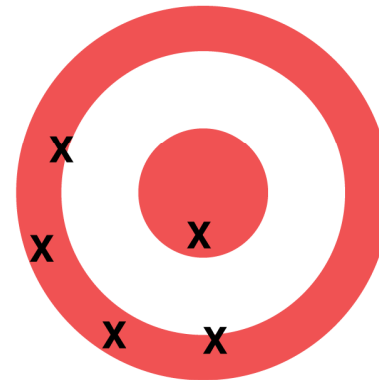
ACCURATE



NOISY



BIASED



BIASED & NOISY

Mean Squared Error

The **mean squared error** of an estimator $\hat{\Theta}$ of the parameter θ is defined as:

$$\text{MSE}(\hat{\Theta}) = E(\hat{\Theta} - \theta)^2 \quad (7-7)$$

Can be rewritten as

$$\begin{aligned} &= E[\hat{\Theta} - E(\hat{\Theta})]^2 + [\theta - E(\hat{\Theta})]^2 \\ &= V(\hat{\Theta}) + (\text{bias})^2 \end{aligned}$$

Statistic #1: Sample Mean

If the values of n observations in a random sample are denoted by x_1, x_2, \dots, x_n , the **sample mean** is

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} \quad (6-1)$$

New random variable \bar{X} is a linear combination of n independent identically distributed variables X_1, X_2, \dots, X_n

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Sample mean \bar{x} is drawn from a random variable

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

$$E(\bar{X}) = \frac{n \cdot E(X_i)}{n} = \frac{n \cdot \mu}{n} = \mu$$

Sample mean, \bar{X} , is an unbiased estimator of the population mean, μ

Sample variance S^2 –
is an estimator of
the population variance σ^2

Sample Variance

If n observations in a sample are denoted by x_1, x_2, \dots, x_n , the **sample variance** is

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (6-3)$$

If one knows the **population average, μ** , one **divides by n** to estimate the variance

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

Why divide by n-1 instead of n?

- The **sample mean \bar{x} is on average closer** to points x_1, x_2, \dots, x_n than **the true mean μ**
$$\sum_{i=1}^n (x_i - \bar{x})^2 \geq \sum_{i=1}^n (x_i - \mu)^2$$
- Consider a sample of size $n=1$.
Then $\bar{x} = x_1$ while $\mu \neq x_1$. Dividing by n gives $s^2 = 0$, while dividing by $n-1$ leaves **s^2 undefined (0/0)**
- For $n=2$, \bar{x} is exactly halfway between x_1 and x_2 making its **sum of squares smaller than** that of μ
- Dividing by $n-1$ on average corrects for a smaller sum of squares: **S^2 is an unbiased estimator of σ^2**

Show that s^2 is unbiased estimate of σ^2

$$\begin{aligned} E(s^2) &= E\left(\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}\right) = \frac{1}{n-1} E\left[\sum_{i=1}^n (X_i^2 + \bar{X}^2 - 2\bar{X}X_i)\right] \\ &= \frac{1}{n-1} E\left[\sum_{i=1}^n X_i^2 + n\bar{X}^2 - 2\bar{X}n\bar{X}\right] = \\ &= \frac{1}{n-1} E\left(\sum_{i=1}^n X_i^2 - n\bar{X}^2\right) = \frac{1}{n-1} (nE(X_i^2) - nE(\bar{X}^2)) \\ &= \frac{1}{n-1} \left(n(\mu^2 + \sigma^2) - n\left(\mu^2 + \frac{\sigma^2}{n}\right)\right) = \frac{n-1}{n-1} \sigma^2 = \underline{\underline{\sigma^2}} \end{aligned}$$

Example 7-4: Sample Variance S^2 is Unbiased

$$\begin{aligned} E(S^2) &= E\left(\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}\right) \\ &= \frac{1}{n-1} E\left[\sum_{i=1}^n (X_i^2 + \bar{X}^2 - 2\bar{X}X_i)\right] \\ &= \frac{1}{n-1} \left[E\left(\sum_{i=1}^n X_i^2 - n\bar{X}^2\right) \right] \\ &= \frac{1}{n-1} \left[\sum_{i=1}^n (\mu^2 + \sigma^2) - n\left(\mu^2 + \frac{\sigma^2}{n}\right) \right] \\ &= \frac{1}{n-1} [n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2] = \frac{1}{n-1} [(n-1)\sigma^2] \end{aligned}$$

Credit: XKCD
comics

WHY ARE THERE SLAVES IN THE BIBLE

WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS

WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT

QUESTIONS FOUND IN GOOGLE AUTOCOMPLETE



WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY

WHY DO WHALES JUMP
WHY ARE WITCHES GREEN
WHY ARE THERE MIRRORS ABOVE BEDS
WHY DO I SAY UH
WHY IS SEA SALT BETTER
WHY ARE THERE TREES IN THE MIDDLE OF FIELDS
WHY IS THERE NOT A POKEMON MMO
WHY IS THERE LAUGHING IN TV SHOWS
WHY ARE THERE DOORS ON THE FREEWAY
WHY ARE THERE SO MANY SVCHOST.EXE RUNNING
WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA
WHY ARE THERE SCARY SOUNDS IN MINECRAFT
WHY IS THERE KICKING IN MY STOMACH
WHY ARE THERE TWO SLASHES AFTER HTTP
WHY ARE THERE CELEBRITIES
WHY DO SNAKES EXIST
WHY DO OYSTERS HAVE PEARLS
WHY ARE DUCKS CALLED DUCKS
WHY DO THEY CALL IT THE CLAP
WHY ARE KYLE AND CARTMAN FRIENDS
WHY IS THERE AN ARROW ON AANG'S HEAD
WHY ARE TEXT MESSAGES BLUE
WHY ARE THERE MUSTACHES ON CLOTHES
WHY ARE THERE MUSTACHES ON CARS
WHY ARE THERE MUSTACHES EVERYWHERE
WHY ARE THERE SO MANY BIRDS IN OHIO
WHY IS THERE SO MUCH RAIN IN OHIO
WHY IS OHIO WEATHER SO WEIRD

WHY AREN'T ECONOMISTS RICH
WHY DO AMERICANS CALL IT SOCCER
WHY ARE MY EARS RINGING
WHY ARE THERE SO MANY AVENGERS
WHY ARE THE AVENGERS FIGHTING THE X MEN
WHY IS WOLVERINE NOT IN THE AVENGERS

WHY ARE THERE SWARMS OF GNATS
WHY IS THERE PHLEGM
WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY IS PSYCHIC WEAK TO BUG
WHY DO CHILDREN GET CANCER
WHY IS POSEIDON ANGRY WITH ODYSSEUS
WHY IS THERE ICE IN SPACE

WHY ARE THERE ANTS IN MY LAPTOP

WHY ARE THERE BRIDESMAIDS
WHY DO DYING PEOPLE REACH UP
WHY AREN'T THERE VARICOSE ARTERIES
WHY ARE OLD KUNGONS DIFFERENT



WHY ARE THERE TINY SPIDERS IN MY HOUSE
WHY DO SPIDERS COME INSIDE
WHY ARE THERE HUGE SPIDERS IN MY HOUSE
WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE
WHY ARE THERE SPIDERS IN MY ROOM
WHY ARE THERE SO MANY SPIDERS IN MY ROOM
WHY DO SPIDER BITES ITCH
WHY IS DYING SO SCARY



WHY IS THERE AN OWL IN MY BACKYARD
WHY IS THERE AN OWL OUTSIDE MY WINDOW
WHY IS THERE AN OWL ON THE DOLLAR BILL
WHY DO OWLS ATTACK PEOPLE
WHY ARE AK 47s SO EXPENSIVE
WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE
WHY ARE THERE GODS
WHY ARE THERE TWO SPOCKS

WHY IS MT VESUVIUS THERE
WHY DO THEY SAY T MINUS
WHY ARE THERE OBELISKS
WHY ARE WRESTLERS ALWAYS WET
WHY ARE OCEANS BECOMING MORE ACIDIC
WHY IS ARWEN DYING
WHY AREN'T MY QUAIL LAYING EGGS
WHY AREN'T MY QUAIL EGGS HATCHING
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA

WHY ARE CIGARETTES LEGAL
WHY ARE THERE DUCKS IN MY POOL
WHY IS JESUS WHITE
WHY IS THERE LIQUID IN MY EAR
WHY DO Q TIPS FEEL GOOD
WHY DO GOOD PEOPLE DIE



WHY ARE DOGS AFRAID OF FIREWORKS
WHY IS THERE NO KING IN ENGLAND

WHY IS PROGRAMMING SO HARD
WHY IS THERE A 0 OHM RESISTOR
WHY DO AMERICANS HATE SOCCER
WHY DO RHYMES SOUND GOOD
WHY DO TREES DIE
WHY IS THERE NO SOUND ON CNN
WHY AREN'T POKEMON REAL
WHY AREN'T BULLETS SHARP
WHY DO DREAMS SEEM SO REAL

WHY IS SEX SO IMPORTANT



WHY IS GPS FREE

WHY IS THERE NO GPS IN LAPTOPS
WHY DO KNEES CLICK
WHY AREN'T THERE E GRADES
WHY IS ISOLATION BAD
WHY DO BOYS LIKE ME
WHY DON'T BOYS LIKE ME
WHY IS THERE ALWAYS A JAVA UPDATE
WHY ARE THERE RED DOTS ON MY THIGHS
WHY IS LYING GOOD

WHY DO IGUANAS DIE
WHY AREN'T THERE DINOSAUR GHOSTS

WHY IS THERE LAVA

WHY ARE THERE ZIPPER

WHY IS LIFE SO BORING

Methods of Point Estimation

- We will cover two popular methodologies to create point estimates of a population parameter.
 - Method of moments
 - Method of maximum likelihood
- Each approach can be used to create estimators with varying degrees of biasedness and relative MSE efficiencies.

Method of moments for point estimation

What are moments?

- The p-th **population moment** of a random variable is the expected value of X^p
 - First moment: $\mu = \int_{-\infty}^{+\infty} x f(x) dx$
 - Second moment: $\mu^2 + \sigma^2 = \int_{-\infty}^{+\infty} x^2 f(x) dx$
 - p-th moment: $\int_{-\infty}^{+\infty} x^p f(x) dx$
 - The **population moment** relates to the entire population
- A **sample moment** is calculated like its population moments but for a finite sample
 - Sample first moment = sample mean = $\frac{1}{n} \sum_{i=1}^n x_i$
 - Sample p-th moment $\frac{1}{n} \sum_{i=1}^n x_i^p$

Moment Estimators

Let X_1, X_2, \dots, X_n be a random sample from either a probability mass function or a probability density function with p unknown parameters $\theta_1, \theta_2, \dots, \theta_p$.

The **moment estimators** $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p$ are found by equating the first p population moments to the first p sample moments and solving the resulting simultaneous equations for the unknown parameters.

Exponential Distribution: Moment Estimator-1st moment

- Suppose that x_1, x_2, \dots, x_n is a random sample from an exponential distribution $f(x) = \lambda \exp(-\lambda x)$ with parameter λ .
- There is only one parameter to estimate, so equating population and sample first moments, we have one equation: $E(X) = \bar{x}$.
- $E(X) = 1/\lambda$ thus
 $\lambda = 1/\bar{x}$ is the 1st moment estimator.

Matlab exercise

- Generate 100,000 exponentially distributed random numbers with $\lambda=3$: $f(x)=\lambda\exp(-\lambda x)$
 - Use `random('Exponential'...)` but read the manual to know how to introduce parameters.
- Get a moment estimate of lambda based on the 1st moment
- Get a moment estimate of lambda based on the 2nd moment
 - Second moment of the exponential distribution is $E(X^2) = E(X)^2 + \text{Var}(X) = 1/\lambda^2 + 1/\lambda^2 = 2/\lambda^2$
- Get a moment estimate of lambda based on the 20th moment
 - Generally, p-th moment of the exponential distribution is $E(X^p) = p!/\lambda^p$

How I solved it

- Stats=100000;
- Y=random('Exponential', 1/3, Stats, 1);
%parametrization in MATLAB is 1/lambda
- 1/mean(Y) %matching the first moment
% ans = 3.0086
- sqrt(2/mean(Y.^2)) %matching the second moment
% ans = 3.0081
- (factorial(20)/mean(Y.^20))^(1./20) %matching the 20th moment

Credit: XKCD
comics

WHY ARE THERE SLAVES IN THE BIBLE

WHY DO TWINS HAVE DIFFERENT FINGERPRINTS
WHY ARE AMERICANS AFRAID OF DRAGONS

WHY IS HTTPS CROSSED OUT IN RED
WHY IS THERE A LINE THROUGH HTTPS
WHY IS THERE A RED LINE THROUGH HTTPS ON FACEBOOK
WHY IS HTTPS IMPORTANT

QUESTIONS FOUND IN GOOGLE AUTOCOMLETE



WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY

WHY DO WHALES JUMP
WHY ARE WITCHES GREEN
WHY ARE THERE MIRRORS ABOVE BEDS
WHY DO I SAY UH
WHY IS SEA SALT BETTER
WHY ARE THERE TREES IN THE MIDDLE OF FIELDS
WHY IS THERE NOT A POKEMON MMO
WHY IS THERE LAUGHING IN TV SHOWS
WHY ARE THERE DOORS ON THE FREEWAY
WHY ARE THERE SO MANY SVCHOST.EXE RUNNING
WHY AREN'T THERE ANY COUNTRIES IN ANTARCTICA
WHY ARE THERE SCARY SOUNDS IN MINECRAFT
WHY IS THERE KICKING IN MY STOMACH
WHY ARE THERE TWO SLASHES AFTER HTTP
WHY ARE THERE CELEBRITIES
WHY DO SNAKES EXIST
WHY DO OYSTERS HAVE PEARLS
WHY ARE DUCKS CALLED DUCKS
WHY DO THEY CALL IT THE CLAP
WHY ARE KYLE AND CARTMAN FRIENDS
WHY IS THERE AN ARROW ON AANG'S HEAD
WHY ARE TEXT MESSAGES BLUE
WHY ARE THERE MUSTACHES ON CLOTHES
WHY ARE THERE MUSTACHES ON CARS
WHY ARE THERE MUSTACHES EVERYWHERE
WHY ARE THERE SO MANY BIRDS IN OHIO
WHY IS THERE SO MUCH RAIN IN OHIO
WHY IS OHIO WEATHER SO WEIRD

WHY AREN'T ECONOMISTS RICH
WHY DO AMERICANS CALL IT SOCCER
WHY ARE MY EARS RINGING
WHY ARE THERE SO MANY AVENGERS
WHY ARE THE AVENGERS FIGHTING THE X MEN
WHY IS WOLVERINE NOT IN THE AVENGERS

WHY ARE THERE SWARMS OF GNATS
WHY IS THERE PHLEGM
WHY ARE THERE SO MANY CROWS IN ROCHESTER, MN
WHY IS PSYCHIC WEAK TO BUG
WHY DO CHILDREN GET CANCER
WHY IS POSEIDON ANGRY WITH ODYSSEUS
WHY IS THERE ICE IN SPACE

WHY ARE THERE ANTS IN MY LAPTOP

WHY ARE THERE BRIDESMAIDS
WHY DO DYING PEOPLE REACH UP
WHY AREN'T THERE VARICOSE ARTERIES
WHY ARE OLD KUNGONS DIFFERENT



WHY ARE THERE TINY SPIDERS IN MY HOUSE
WHY DO SPIDERS COME INSIDE
WHY ARE THERE HUGE SPIDERS IN MY HOUSE
WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE
WHY ARE THERE SPIDERS IN MY ROOM
WHY ARE THERE SO MANY SPIDERS IN MY ROOM
WHY DO SPIDER BITES ITCH
WHY IS DYING SO SCARY



WHY IS THERE AN OWL IN MY BACKYARD
WHY IS THERE AN OWL OUTSIDE MY WINDOW
WHY IS THERE AN OWL ON THE DOLLAR BILL
WHY DO OWLS ATTACK PEOPLE
WHY ARE AK 47s SO EXPENSIVE
WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE
WHY ARE THERE GODS
WHY ARE THERE TWO SPOCKS

WHY ARE DOGS AFRAID OF FIREWORKS
WHY IS THERE NO KING IN ENGLAND

WHY IS PROGRAMMING SO HARD
WHY IS THERE A 0 OHM RESISTOR
WHY DO AMERICANS HATE SOCCER
WHY DO RHYMES SOUND GOOD
WHY DO TREES DIE
WHY IS THERE NO SOUND ON CNN
WHY AREN'T POKEMON REAL
WHY AREN'T BULLETS SHARP
WHY DO DREAMS SEEM SO REAL

WHY IS THERE NO GPS IN LAPTOPS
WHY DO KNEES CLICK
WHY AREN'T THERE E GRADES
WHY IS ISOLATION BAD
WHY DO BOYS LIKE ME
WHY DON'T BOYS LIKE ME
WHY IS THERE ALWAYS A JAVA UPDATE
WHY ARE THERE RED DOTS ON MY THIGHS
WHY IS LYING GOOD



WHY IS MT VESUVIUS THERE
WHY DO THEY SAY T MINUS
WHY ARE THERE OBELISKS
WHY ARE WRESTLERS ALWAYS WET
WHY ARE OCEANS BECOMING MORE ACIDIC
WHY IS ARWEN DYING
WHY AREN'T MY QUAIL LAYING EGGS
WHY AREN'T MY QUAIL EGGS HATCHING
WHY AREN'T THERE ANY FOREIGN MILITARY BASES IN AMERICA

WHY ARE CIGARETTES LEGAL
WHY ARE THERE DUCKS IN MY POOL
WHY IS JESUS WHITE
WHY IS THERE LIQUID IN MY EAR
WHY DO Q TIPS FEEL GOOD
WHY DO GOOD PEOPLE DIE



WHY ARE ULTRASOUNDS IMPORTANT
WHY ARE ULTRASOUND MACHINES EXPENSIVE
WHY IS STEALING WRONG

Method of Maximum Likelihood for point estimation

Maximum Likelihood Estimators

- Suppose that X is a random variable with probability distribution $f(x, \theta)$, where θ is a single unknown parameter. Let x_1, x_2, \dots, x_n be the observed values in a random sample of size n . Then the **likelihood function** of the sample is the probability to get it in a random variable with PDF $f(x, \theta)$:

$$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdot \dots \cdot f(x_n, \theta) \quad (7-9)$$

- Note that the likelihood function is now a function of only the unknown parameter θ . The **maximum likelihood estimator** (MLE) of θ is the value of θ that maximizes the likelihood function $L(\theta)$.
- Usually, it is easier to work with **logarithms**: $l(\theta) = \ln L(\theta)$

Exponential MLE:

$$f(x_i) = \lambda e^{-\lambda x_i}$$

$$L(\lambda) = P(x_1, x_2, \dots, x_n | \lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} =$$

$$= \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$$

$$\ln L(\lambda) = n \ln(\lambda) - \lambda \sum_{i=1}^n x_i$$

$$\frac{d \ln L(\lambda)}{d \lambda} = \frac{n}{\lambda} - \sum x_i = 0$$

$$\hat{\lambda} = \frac{n}{\sum x_i} = \frac{1}{\bar{X}}$$

Same as
1st moment
estimator

Example 7-11: Exponential MLE

Let X be an exponential random variable with parameter λ . The likelihood function of a random sample of size n is:

$$L(\lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$$

$$\ln L(\lambda) = n \ln(\lambda) - \lambda \sum_{i=1}^n x_i$$

$$\frac{d \ln L(\lambda)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0$$

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{X}} \quad (\text{same as moment estimator})$$

Bernoulli; MLE

$$f(x, p) = p^x (1-p)^{1-x}$$

$$L(p) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} =$$

$$= p^{\sum x_i} (1-p)^{n - \sum x_i}$$

$$\ln L(p) = (\sum x_i) \ln p + (n - \sum x_i) \ln(1-p)$$

$$\frac{d \ln L(p)}{dp} = \frac{\sum x_i}{p} - \frac{n - \sum x_i}{1-p} = 0 \quad \text{at } \hat{p}$$

$$0 = \frac{(1 - \hat{p}) \sum x_i - \hat{p} (n - \sum x_i)}{\hat{p} (1 - \hat{p})} \quad \hat{p} = \frac{\sum_{i=1}^n x_i}{n}$$

Example 7-9: Bernoulli MLE

Let X be a Bernoulli random variable. The probability mass function is $f(x;p) = p^x(1-p)^{1-x}$, $x = 0, 1$ where P is the parameter to be estimated. The likelihood function of a random sample of size n is:

$$\begin{aligned} L(p) &= p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \cdot \dots \cdot p^{x_n}(1-p)^{1-x_n} \\ &= \prod_{i=1}^n p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i} \end{aligned}$$

$$\ln L(p) = \left(\sum_{i=1}^n x_i \right) \ln p + \left(n - \sum_{i=1}^n x_i \right) \ln(1-p)$$

$$\frac{d \ln L(p)}{dp} = \frac{\sum_{i=1}^n x_i}{p} - \frac{(n - \sum_{i=1}^n x_i)}{(1-p)} = 0$$

$$\hat{p} = \frac{\sum_{i=1}^n x_i}{n} \text{ (same as moment estimator)}$$

Normal MLE for μ

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$L(\mu, \sigma) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{\sum (x_i - \mu)^2}{2\sigma^2}\right)$$

$$\ln L(\mu, \sigma) = -n \ln(\sigma\sqrt{2\pi}) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

$$\frac{d \ln L(\mu, \sigma)}{d\mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0 \text{ at } \hat{\mu}$$
$$\hat{\mu} = \frac{\sum_{i=1}^n x_i}{n}$$

Example 7-10: Normal MLE for μ

Let X be a normal random variable with unknown mean μ and variance σ^2 . The likelihood function of a random sample of size n is:

$$\begin{aligned}L(\mu) &= \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\&= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i-\mu)^2} \\ \ln L(\mu) &= \frac{-n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \\ \frac{d \ln L(\mu)}{d\mu} &= \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0 \\ \hat{\mu} &= \frac{\sum_{i=1}^n x_i}{n} = \bar{X} \text{ (same as moment estimator)}\end{aligned}$$

Example 7-11: Normal MLE for σ^2

Let X be a normal random variable with the estimate of mean μ determined by MLE (see the previous slide) and an **unknown variance σ^2** . The likelihood function of a random sample of size n is:

$$\begin{aligned}L(\sigma) &= \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i-\mu)^2}{(2\sigma^2)}} \\&= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{\frac{-1}{2\sigma^2} \sum_{i=1}^n (x_i-\mu)^2} \\ \ln L(\sigma) &= \frac{-n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \\ \frac{d \ln L(\sigma)}{d\sigma} &= -\frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^n (x_i - \mu)^2 = 0 \\ \widehat{\sigma^2} &= \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \quad (\text{biased estimator})\end{aligned}$$