

# 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  $\lambda$
  - For Bernoulli PDF:  $p^x(1-p)^{1-x}$  one wants to estimate  $p$
  - For normal PDF one wants to estimate both  $\mu$  and  $\sigma$
- Point estimates are uncertain: therefore, we can talk of averages and standard deviations of point estimators

# 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**,  $\mu$ , 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  $\mu$**   
$$\sum_{i=1}^n (x_i - \bar{x})^2 \geq \sum_{i=1}^n (x_i - \mu)^2$$
- 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  $\mu$
- Dividing by  $n-1$  on average corrects for a smaller sum of squares:  **$S^2$  is an unbiased estimator of  $\sigma^2$**



Show that  $s^2$  is unbiased estimate of  $\sigma^2$

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

## 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 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 ARE THERE GHOSTS

WHY IS THERE AN OWL IN MY BACKYARD

WHY IS SPACE BLACK

WHY ARE THERE GHOSTS

WHY IS THERE AN OWL OUTSIDE MY WINDOW

WHY IS OUTER SPACE SO COLD

WHY ARE THERE GHOSTS

WHY IS THERE AN OWL ON THE DOLLAR BILL

WHY ARE THERE PYRAMIDS ON THE MOON

WHY ARE THERE GHOSTS

WHY DO OWLS ATTACK PEOPLE

WHY IS NASA SHUTTING DOWN

WHY ARE THERE GHOSTS

WHY ARE AK 47s SO EXPENSIVE

WHY ARE THERE MALE AND FEMALE BIKES

WHY ARE THERE GHOSTS

WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE

WHY ARE THERE TINY SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY ARE THERE GODS

WHY DO SPIDERS COME INSIDE

WHY ARE THERE GHOSTS

WHY ARE THERE TWO SPOCKS

WHY ARE THERE HUGE SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY IS LIFE SO BORING

WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY ARE CIGARETTES LEGAL

WHY ARE THERE SPIDERS IN MY ROOM

WHY ARE THERE GHOSTS

WHY ARE THERE DUCKS IN MY POOL

WHY ARE THERE SO MANY SPIDERS IN MY ROOM

WHY ARE THERE GHOSTS

WHY IS JESUS WHITE

WHY DO SPIDER BITES ITCH

WHY ARE THERE GHOSTS

WHY IS THERE LIQUID IN MY EAR

WHY IS DYING SO SCARY

WHY ARE THERE GHOSTS

WHY DO Q TIPS FEEL GOOD

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

WHY AREN'T THERE DINOSAUR GHOSTS

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 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 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 ARE THERE GHOSTS

WHY IS THERE AN OWL IN MY BACKYARD

WHY IS SPACE BLACK

WHY ARE THERE GHOSTS

WHY IS THERE AN OWL OUTSIDE MY WINDOW

WHY IS OUTER SPACE SO COLD

WHY ARE THERE GHOSTS

WHY IS THERE AN OWL ON THE DOLLAR BILL

WHY ARE THERE PYRAMIDS ON THE MOON

WHY ARE THERE GHOSTS

WHY DO OWLS ATTACK PEOPLE

WHY IS NASA SHUTTING DOWN

WHY ARE THERE GHOSTS

WHY ARE AK 47s SO EXPENSIVE

WHY ARE THERE MALE AND FEMALE BIKES

WHY ARE THERE GHOSTS

WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE

WHY ARE THERE TINY SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY ARE THERE GODS

WHY DO SPIDERS COME INSIDE

WHY ARE THERE GHOSTS

WHY ARE THERE TWO SPOCKS

WHY ARE THERE HUGE SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY IS LIFE SO BORING

WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY ARE CIGARETTES LEGAL

WHY ARE THERE SPIDERS IN MY ROOM

WHY ARE THERE GHOSTS

WHY ARE THERE DUCKS IN MY POOL

WHY ARE THERE SO MANY SPIDERS IN MY ROOM

WHY ARE THERE GHOSTS

WHY IS JESUS WHITE

WHY DO SPIDER BITES ITCH

WHY ARE THERE GHOSTS

WHY IS THERE LIQUID IN MY EAR

WHY IS DYING SO SCARY

WHY ARE THERE GHOSTS

WHY DO Q TIPS FEEL GOOD



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

WHY IS SEX SO IMPORTANT



WHY ARE THERE SQUIRRELS

# 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-1<sup>st</sup> 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  $\lambda$ .
- 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 1<sup>st</sup> moment estimator.

# Matlab exercise

- Generate  $n=100$  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 1<sup>st</sup> moment
- Get a moment estimate of lambda based on the 2<sup>nd</sup> 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 each of the first 30 moments and plot relative error vs moment  $p$ 
  - Generally,  $p$ -th moment of the exponential distribution is  $E(X^p) = p!/\lambda^p$

# How I solved it

- `n=100;`
- `lambda_true=3;`
- `Y = random('Exponential',1./lambda_true, n, 1);`
- `%%`
- `% first moment estimator based on mean(Y)`
- `lambda_1_hat=n./sum(Y);`
- `%%`
- `% second moment estimator based on mean(Y.^2)=2./lambda.^2`
- `sample_moment_2=sum(Y.^2)./n;`
- `lambda_2_hat=(2./sample_moment_2).^0.5 %matching the second moment`
- `%%`
  
- `for p=1:30;`
- `sample_moment_p=sum(Y.^p)./n;`
- `lambda_p_hat(p)=(factorial(p)/sample_moment_p).^(1./p);`
- `MSE(p)=(lambda_p_hat(p)-lambda_true).^2;`
- `end`
- `% estimate the fraction of error in p-th moment`
- `% parameter estimator as a function of p`
- `figure; semilogy(sqrt(MSE)./lambda_true,'ko-');`
- `%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 DO IGUANAS DIE  
WHY AREN'T THERE DINOSAUR GHOSTS

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

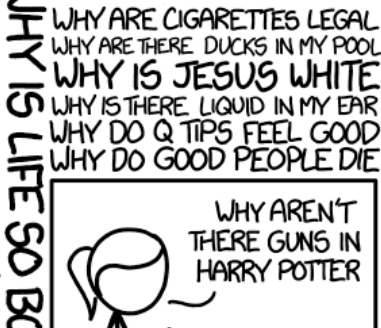
WHY ARE DOGS AFRAID OF FIREWORKS  
WHY IS THERE NO KING IN ENGLAND

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



# Exponential MLF:

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

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

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

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

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

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

Same as  
1<sup>st</sup> moment  
estimator

# Example 7-11: Exponential MLE

Let  $X$  be an exponential random variable with parameter  $\lambda$ . 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  $\mu$

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

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

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

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

# Example 7-10: Normal MLE for $\mu$

Let  $X$  be a normal random variable with unknown mean  $\mu$  and variance  $\sigma^2$ . 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 $\sigma^2$

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

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



# MLE for Poisson distribution

$$\begin{aligned} f(x_1, \dots, x_n | \lambda) &= \frac{e^{-\lambda} \lambda^{x_1}}{x_1!} \dots \frac{e^{-\lambda} \lambda^{x_n}}{x_n!} \\ &= \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i}}{x_1! \dots x_n!} \end{aligned}$$

$$\log f(x_1, \dots, x_n | \lambda) = -n\lambda + \sum_{i=1}^n x_i \log \lambda - \log c$$

where  $c = \prod_{i=1}^n x_i!$  does not depend on  $\lambda$ , and

$$\frac{d}{d\lambda} \log f(x_1, \dots, x_n | \lambda) = -n + \frac{\sum_{i=1}^n x_i}{\lambda}$$

By equating to zero, we obtain that the maximum likelihood estimate  $\hat{\lambda}$  equals

$$\hat{\lambda} = \frac{\sum_{i=1}^n x_i}{n}$$

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 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 ARE THERE GHOSTS

WHY IS THERE AN OWL IN MY BACKYARD

WHY IS SPACE BLACK

WHY ARE THERE GHOSTS

WHY IS THERE AN OWL OUTSIDE MY WINDOW

WHY IS OUTER SPACE SO COLD

WHY ARE THERE GHOSTS

WHY IS THERE AN OWL ON THE DOLLAR BILL

WHY ARE THERE PYRAMIDS ON THE MOON

WHY ARE THERE GHOSTS

WHY DO OWLS ATTACK PEOPLE

WHY IS NASA SHUTTING DOWN

WHY ARE THERE GHOSTS

WHY ARE AK 47s SO EXPENSIVE

WHY ARE THERE MALE AND FEMALE BIKES

WHY ARE THERE GHOSTS

WHY ARE THERE HELICOPTERS CIRCLING MY HOUSE

WHY ARE THERE TINY SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY ARE THERE GODS

WHY DO SPIDERS COME INSIDE

WHY ARE THERE GHOSTS

WHY ARE THERE TWO SPOCKS

WHY ARE THERE HUGE SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY IS LIFE SO BORING

WHY ARE THERE LOTS OF SPIDERS IN MY HOUSE

WHY ARE THERE GHOSTS

WHY ARE CIGARETTES LEGAL

WHY ARE THERE SPIDERS IN MY ROOM

WHY ARE THERE GHOSTS

WHY ARE THERE DUCKS IN MY POOL

WHY ARE THERE SO MANY SPIDERS IN MY ROOM

WHY ARE THERE GHOSTS

WHY IS JESUS WHITE

WHY DO SPIDER BITES ITCH

WHY ARE THERE GHOSTS

WHY IS THERE LIQUID IN MY EAR

WHY IS DYING SO SCARY

WHY ARE THERE GHOSTS

WHY DO Q TIPS FEEL GOOD

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

WHY AREN'T THERE DINOSAUR GHOSTS

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

WHY DO IGUANAS DIE

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

WHY AREN'T THERE DINOSAUR GHOSTS

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

WHY AREN'T THERE DINOSAUR GHOSTS

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

WHY AREN'T THERE DINOSAUR GHOSTS

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

WHY AREN'T THERE DINOSAUR GHOSTS

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

WHY AREN'T THERE DINOSAUR GHOSTS

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

WHY AREN'T THERE DINOSAUR GHOSTS

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

WHY AREN'T THERE DINOSAUR GHOSTS

WHY IS OHIO WEATHER SO WEIRD  
WHY ARE THERE MALE AND FEMALE BIKES  
WHY ARE THERE BRIDESMAIDS

WHY AREN'T THERE DINOSAUR GHOSTS

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

WHY AREN'T THERE DINOSAUR GHOSTS

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

WHY AREN'T THERE DINOSAUR GHOSTS

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

WHY AREN'T THERE DINOSAUR GHOSTS

WHY IS SEX SO IMPORTANT



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 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

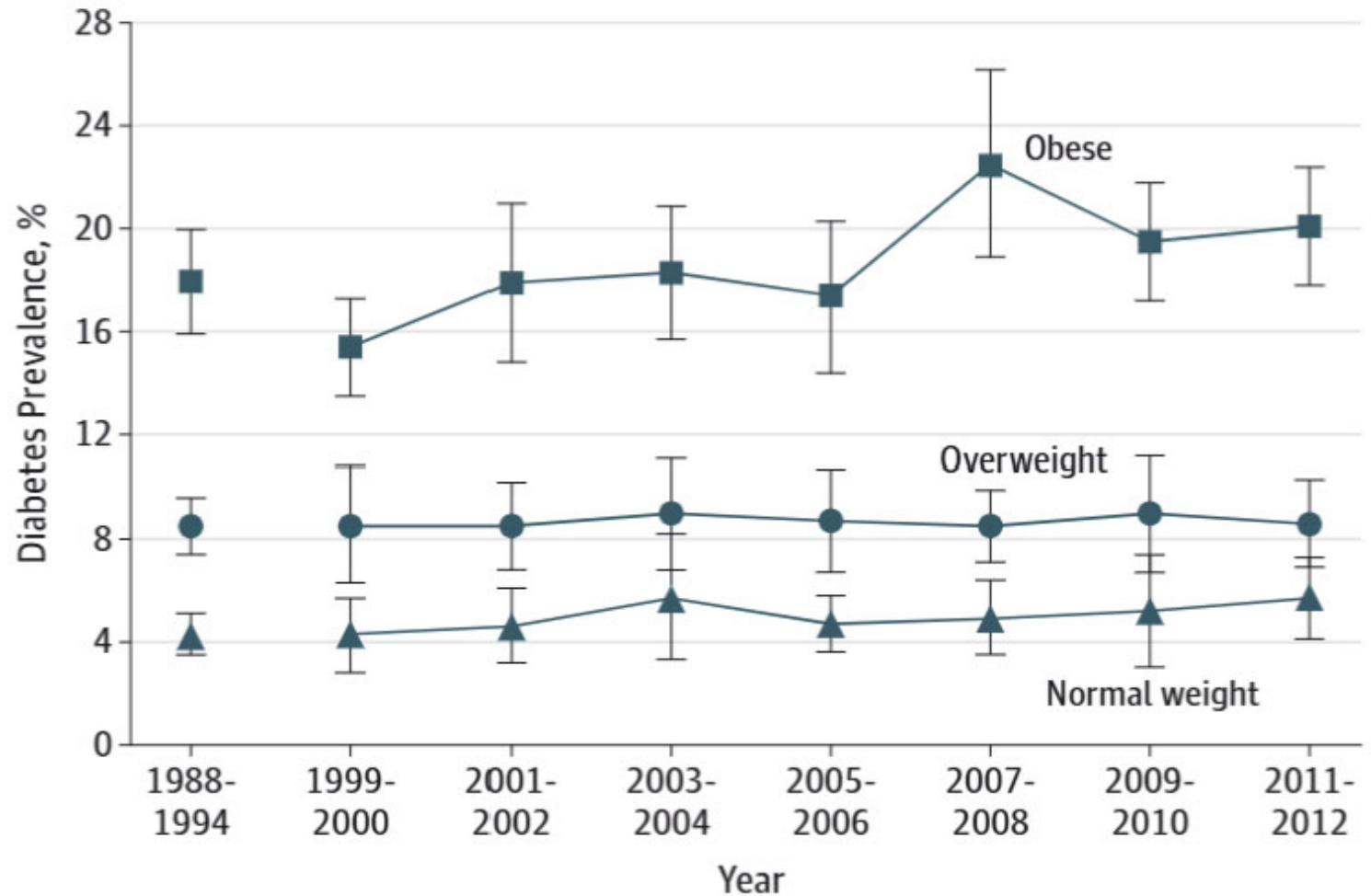


WHY ARE DOGS AFRAID OF FIREWORKS  
WHY IS THERE NO KING IN ENGLAND

# Confidence Intervals

- **Point estimates:** give the best single estimate of a parameter from sample data
- **Confidence intervals** (or interval estimate): give the upper and lower bounds within which a population parameter is with a given confidence level  $1-\alpha$ , say 95% ( $\alpha=0.05$ )

Figure 2. US Trends in Diabetes Prevalence per 100 Adults Aged 20 Years or Older by BMI Category



No. of participants	1988-1994	1999-2000	2001-2002	2003-2004	2005-2006	2007-2008	2009-2010	2011-2012
Obese	2324	727	732	815	820	1137	1302	1075
Overweight	2942	724	878	784	694	949	1009	852
Normal weight	3025	645	699	624	604	726	762	785

# Two-sided confidence intervals

- Calculated based on the sample  $X_1, X_2, \dots, X_n$
- Characterized by:
  - lower- and upper- confidence limits  $L$  and  $R$
  - the confidence coefficient  $1-\alpha$
- Objective: for two-sided confidence interval, find  $L$  and  $R$  such that
  - $\text{Prob}(\mu > R) = \alpha/2$
  - $\text{Prob}(\mu < L) = \alpha/2$
  - Therefore,  $\text{Prob}(L < \mu < R) = 1-\alpha$
- For one-sided confidence interval, say, upper bound of  $\mu$ , find  $R$  that
  - $\text{Prob}(\mu > R) = \alpha$
- **Assume standard deviation sigma is known**



Consider  $1 - \alpha = 95\% = 0.95$

$$\alpha = 0.05; \quad \frac{\alpha}{2} = 0.025$$

$$z_{\alpha/2} = 1.96 \rightarrow \text{Prob}(Z > z_{\alpha/2}) = \frac{\alpha}{2}$$



$$\text{Prob}\left(-z_{\frac{\alpha}{2}} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$\text{Prob}\left(\bar{X} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

For one sided lower bound on  $\mu$

$$\text{Prob}\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \underline{z_{\alpha}}\right) \rightarrow$$

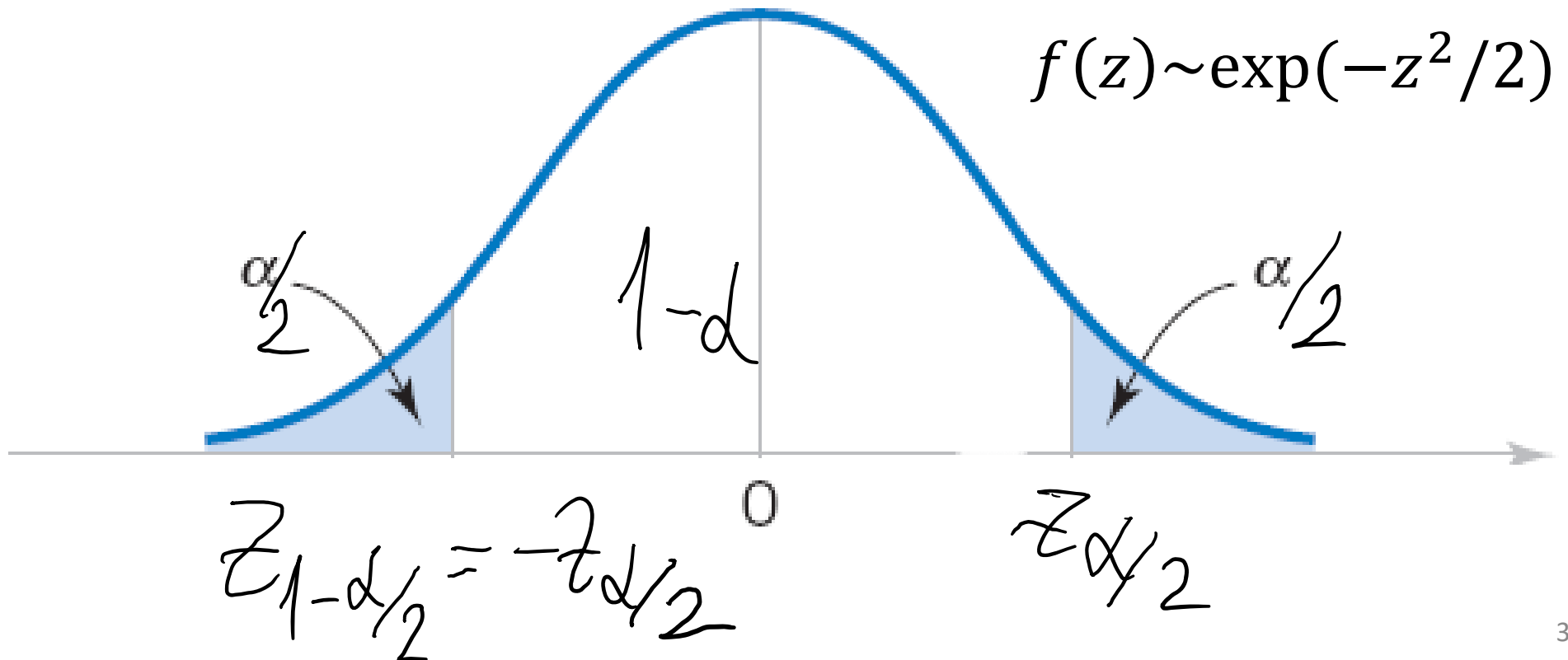
$$\mu > \bar{X} - z_{\alpha} \frac{\sigma}{\sqrt{n}}$$

$$z_{\alpha} = 1.65 < \\ z_{\alpha/2} = 1.96$$

# Confidence Interval on the Population Mean, Variance Known

---

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$



# Exercise

Ishikawa et al. (Journal of Bioscience and Bioengineering 2012) studied the force with which bacterial biofilms adhere to a solid surface.

Five measurements for a bacterial strain of *Acinetobacter* gave readings 2.69, 5.76, 2.67, 1.62, and 4.12 dyne-cm<sup>2</sup>.

Assume that the standard deviation is known to be 0.66 dyne-cm<sup>2</sup>

- (a) Find 95% confidence interval for the mean adhesion force
- (b) If scientists want the width of the confidence interval to be below 0.55 dyne-cm<sup>2</sup> what number of samples should be?

Ishikawa et al. (Journal of Bioscience and Bioengineering 2012) studied the force with which bacterial biofilms adhere to a solid surface. Five measurements for a bacterial strain of Acinetobacter gave readings 2.69, 5.76, 2.67, 1.62, and 4.12 dyne-cm<sup>2</sup>. Assume that the **standard deviation is known to be 0.66 dyne-cm<sup>2</sup>**

- (a) Find 95% confidence interval for the mean adhesion force
- (b) If scientists want the width of the confidence interval to be below 0.55 dyne-cm<sup>2</sup> what number of samples should be?

a) 95% CI for  $\mu$ ,  $n = 5$   $\sigma = 0.66$   $\bar{x} = 3.372$ ,  $z = 1.96$

$$\bar{x} - z\sigma / \sqrt{n} \leq \mu \leq \bar{x} + z\sigma / \sqrt{n}$$

$$3.372 - 1.96(0.66 / \sqrt{5}) \leq \mu \leq 3.372 + 1.96(0.66 / \sqrt{5})$$

$$2.79 \leq \mu \leq 3.95$$

b) Width is  $2z\sigma / \sqrt{n} = 0.55$ , therefore  $n = [2z\sigma / 0.55]^2 = [2(1.96)(0.66) / 0.55]^2 = 22.13$   
Round up to  $n = 23$ .