

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

WHY AREN'T MY
ARMS GROWING



WHY ARE THERE WEEKS
WHY DO I FEEL DIZZY

QUESTIONS

FOUND IN GOOGLE AUTOCOMPLETE

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 ANTS IN MY LAPTOP

WHY IS EARTH TILTED

WHY IS SPACE BLACK

WHY IS OUTER SPACE SO COLD

WHY ARE THERE PYRAMIDS ON THE MOON

WHY IS NASA SHUTTING DOWN

WHY ARE THERE MALE AND FEMALE BIKES

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 NO GPS IN LAPTOPS

WHY DO KNEES CLICK

WHY AREN'T THERE E GRADES

WHY IS SEX
SO IMPORTANT



WHY ARE THERE
GHOSTS



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 AREN'T
THERE GUNS IN
HARRY POTTER



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

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 KINGDOMS 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 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 HELL IF GOD FORGIVES
WHY IS GPS FREE

WHY ARE THERE FEMALE MR NIMES

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 DOGS AFRAID OF FIREWORKS

WHY IS THERE NO KING IN ENGLAND

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 DOGS AFRAID OF FIREWORKS

WHY IS THERE NO KING IN ENGLAND

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 DOGS AFRAID OF FIREWORKS

WHY IS THERE NO KING IN ENGLAND

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 DOGS AFRAID OF FIREWORKS

WHY IS THERE NO KING IN ENGLAND

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 DOGS AFRAID OF FIREWORKS

WHY IS THERE NO KING IN ENGLAND

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 DOGS AFRAID OF FIREWORKS

WHY IS THERE NO KING IN ENGLAND

Mean & Variance of the Sample Average

If $\bar{X} = \frac{(X_1 + X_2 + \dots + X_n)}{n}$ and $E(X_i) = \mu$

$$\text{Then } E(\bar{X}) = \frac{n \cdot \mu}{n} = \mu \quad (5-28a)$$

If the X_i are independent with $V(X_i) = \sigma^2$

$$\text{Then } V(\bar{X}) = \frac{n \cdot \sigma^2}{n^2} = \frac{\sigma^2}{n} \quad (5-28b)$$

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

Test CLT for your own random variable

- Go to:
https://onlinestatbook.com/stat_sim/sampling_dist/
- Select “Custom” at the top and use mouse to sketch the PMF of your own random variable
- Select “mean” and $n=5$ in the third panel
- Choose “Animated” in the second panel and use `number_of_experiments=5` to see one sample being generated
- Repeat with `number_of_experiments = 10,000`
- Now select “mean” and $n=25$ in the fourth panel
- Skewness and Kurtosis are measures of how good is the normal (Gaussian) fit (choose “fit normal”)

Sampling Distributions of Sample Means

Figure 7-1 Distributions of average scores from throwing dice.

Mean = $(6+1)/2=3.5$

$\text{Sigma}^2 = [(6-1+1)^2-1]/12=2.92$

$\text{Sigma}=1.71$

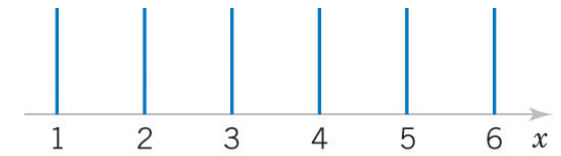
Formulas

$$\mu = \frac{b + a}{2}$$

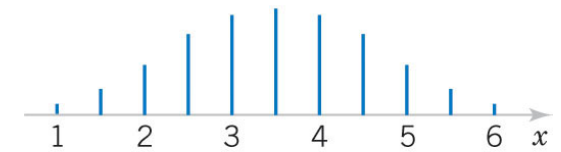
$$\sigma_X^2 = \frac{(b - a + 1)^2 - 1}{12}$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma_X^2}{n}$$

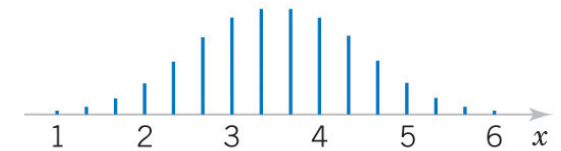
show
Matlab



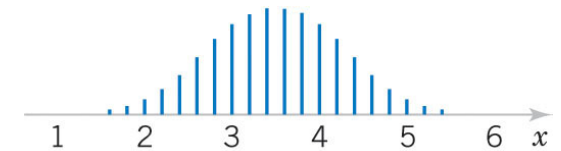
(a) One die



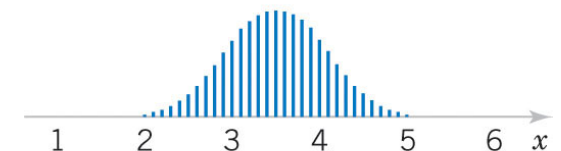
(b) Two dice



(c) Three dice



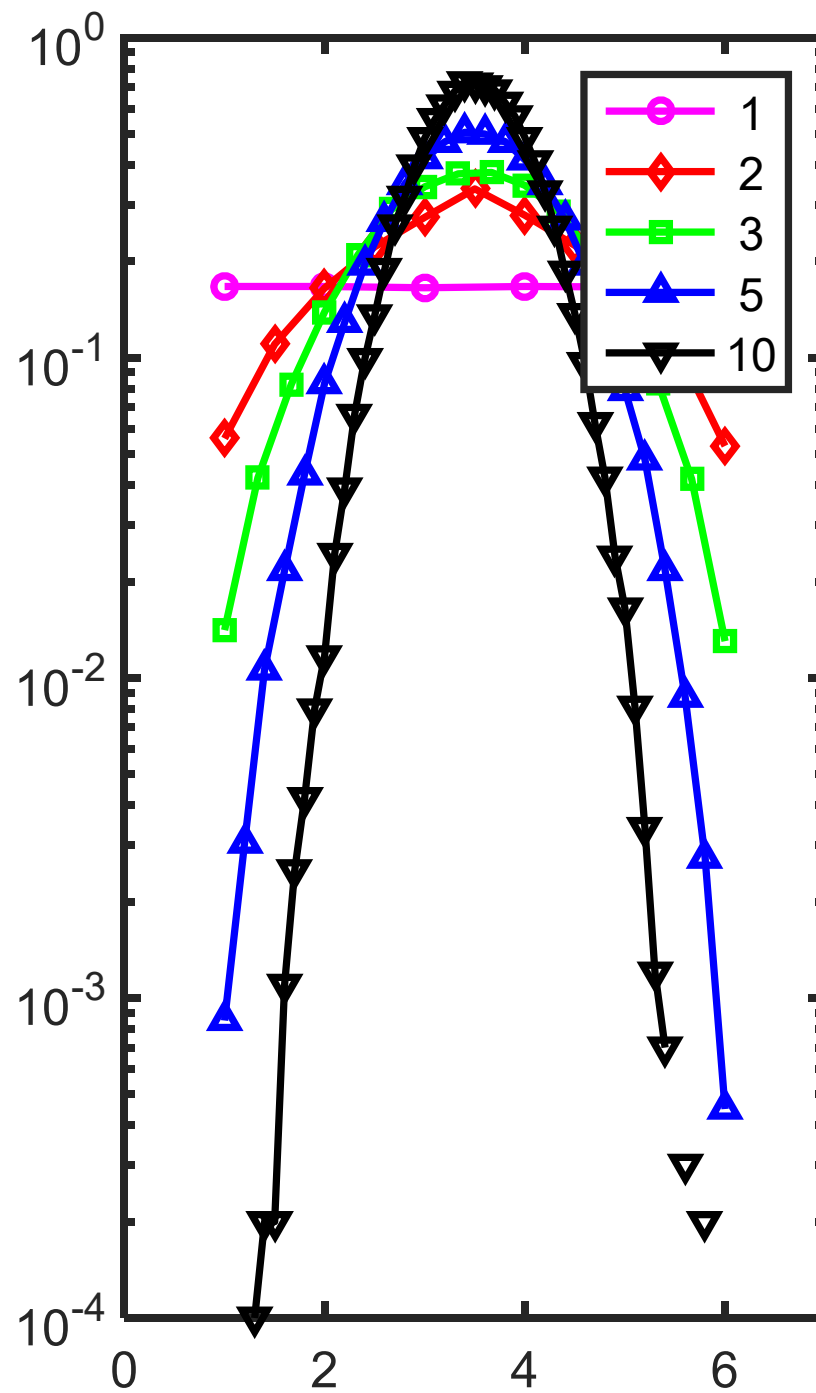
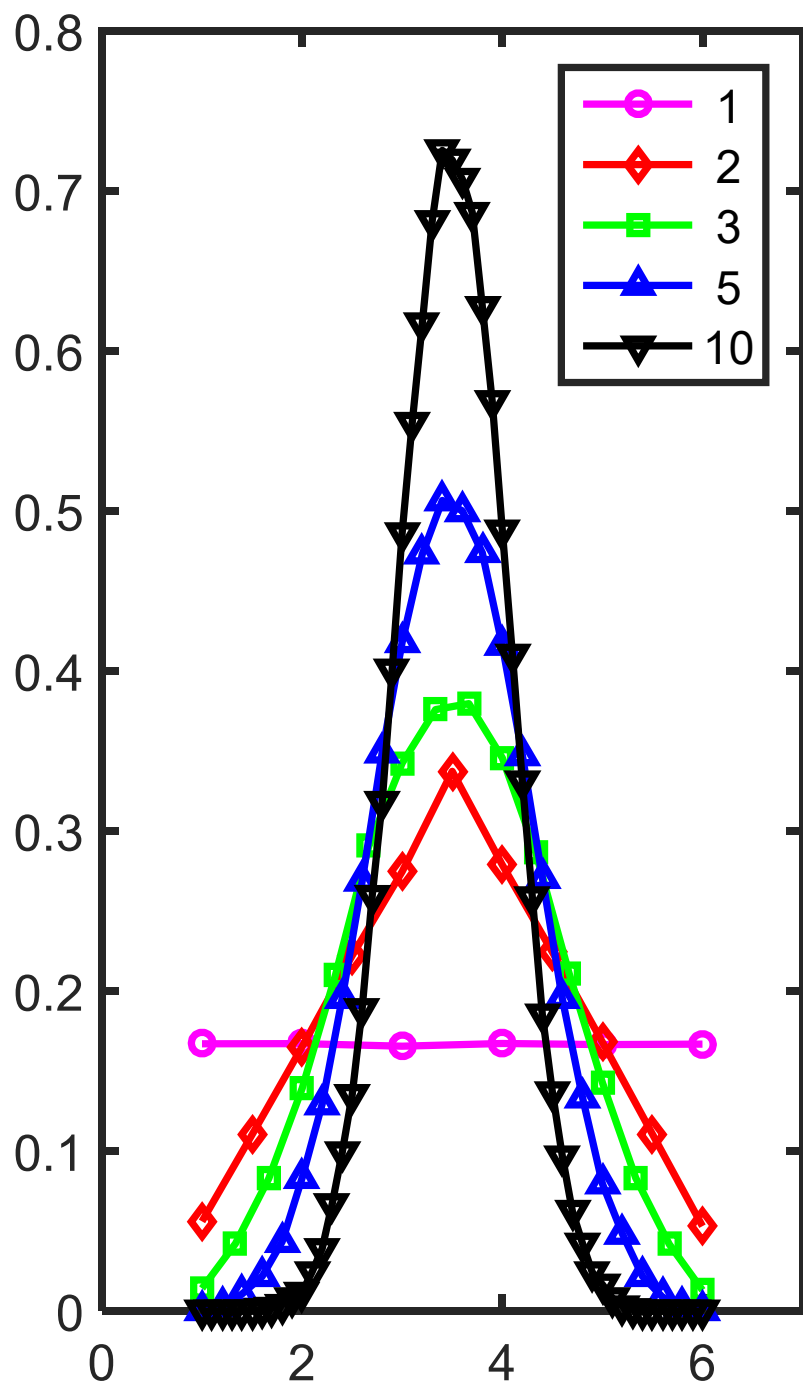
(d) Five dice



(e) Ten dice

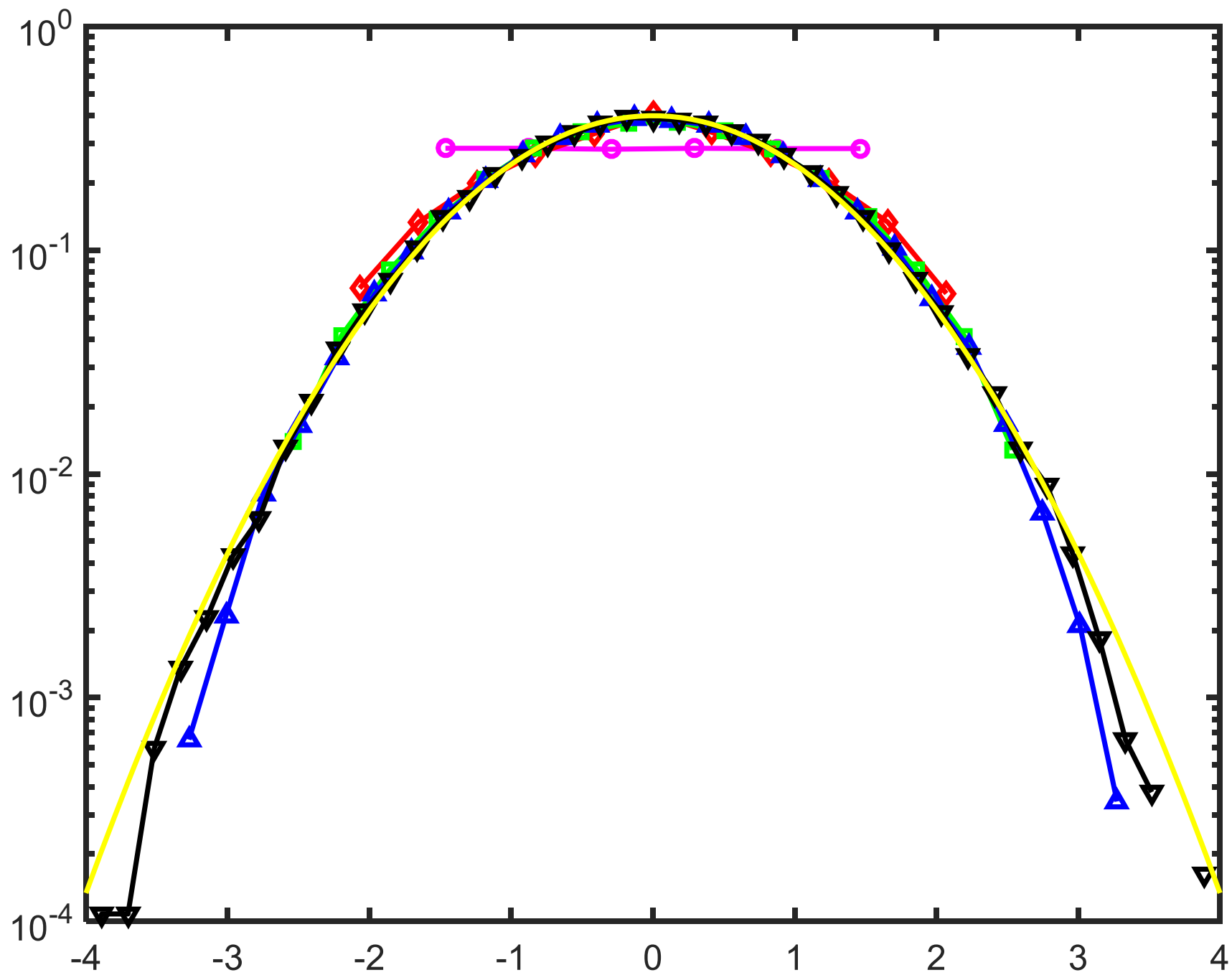
Matlab demonstration

- `Stats=100000; N=10;`
- `r_table=floor(6.*rand(Stats,N))+1;`
- `%%`
- `r1=r_table(:,1);`
- `step=1; [a,b1]=hist(r1,1:step:6);`
- `pdf_r1=a./sum(a)./step;`
- `figure; hold on; subplot(1,2,1); plot(b1,pdf_r1,'mo-'); hold on; axis([0 7 0 0.2]); subplot(1,2,2);`
`semilogy(b1,pdf_r1,'mo-'); hold on; axis([0 7 1e-3 1]);`
- `%%`
- `r2=(r_table(:,1)+r_table(:,2))./2;`
- `step=0.5; [a,b2]=hist(r2,1:step:6); pdf_r2=a./sum(a)./step;`
- `subplot(1,2,1); plot(b2,pdf_r2,'rd-'); axis([0 7 0 0.4]); subplot(1,2,2); semilogy(b2,pdf_r2,'rd-');`
- `%%`
- `r3=(r_table(:,1)+r_table(:,2)+r_table(:,3))./3;`
- `step=1./3; [a,b3]=hist(r3,1:step:6); pdf_r3=a./sum(a)./step;`
- `subplot(1,2,1); plot(b3,pdf_r3,'gs-'); axis([0 7 0 0.4]); subplot(1,2,2); semilogy(b3,pdf_r3,'gs-');`
- `%%`
- `r5=sum(r_table(:,1:5),2)./5;`
- `step=1./5; [a,b5]=hist(r5,1:step:6); pdf_r5=a./sum(a)./step;`
- `subplot(1,2,1); plot(b5,pdf_r5,'b^-'); axis([0 7 0 0.6]); subplot(1,2,2); semilogy(b5,pdf_r5,'b^-'); axis([0 7 1e-4 1]);`
- `%%`
- `r10=sum(r_table(:,1:10),2)./10;`
- `step=1./10; [a,b10]=hist(r10,1:step:6); pdf_r10=a./sum(a)./step;`
- `subplot(1,2,1); plot(b10,pdf_r10,'kv-'); axis([0 7 0 0.8]); legend(num2str([1,2,3,5,10]'));`
- `subplot(1,2,2); semilogy(b10,pdf_r10,'kv-'); legend(num2str([1,2,3,5,10]'));`



Matlab demonstration; part 2

- `%%Now plot all of them normalized to 0 and std 1`
- `sigma=sqrt(35/12);`
- `mu=3.5;`
- `figure;`
- `sigma1=sigma;`
- `semilogy((b1-mu)./sigma1,pdf_r1.*sigma1,'mo-');`
- `axis([-4 4 1e-3 1]);`
- `hold on;`
- `%%`
- `sigma2=sigma./sqrt(2);`
- `semilogy((b2-mu)./sigma2,pdf_r2.*sigma2,'rd-');`
- `%%`
- `sigma3=sigma./sqrt(3);`
- `semilogy((b3-mu)./sigma3,pdf_r3.*sigma3,'gs-');`
- `%%`
- `sigma5=sigma./sqrt(5);`
- `semilogy((b5-mu)./sigma5,pdf_r5.*sigma5,'b^-');`
- `axis([-4 4 1e-4 1]);`
- `%%`
- `sigma10=sigma./sqrt(10);`
- `semilogy((b10-mu)./sigma10,pdf_r10.*sigma10,'kv-');`
- `axis([-4 4 1e-4 1]);`
- `%%`
- `%Let's see how well does the Gaussian fits it`
- `x=-4:0.1:4;`
- `semilogy(x,1./sqrt(2*pi)*exp(-x.^2./2),'y-');`



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?

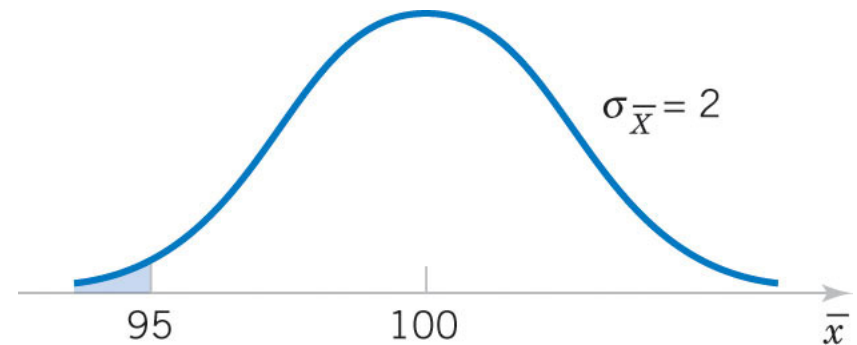


Figure 7-2 Desired probability is shaded

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 $\mu_{\text{old}}=5000$ 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?

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

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

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

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

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

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

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

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?

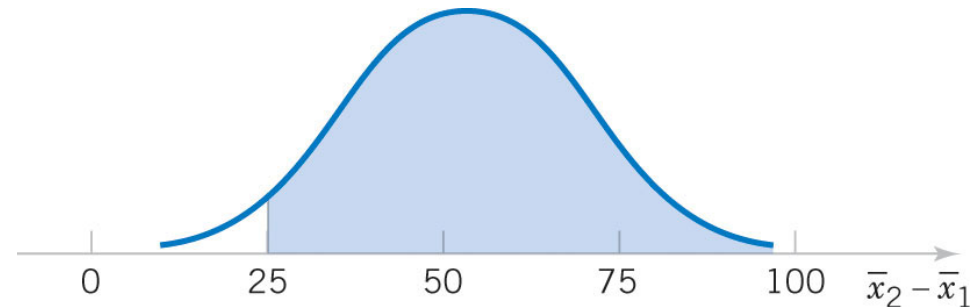


Figure 7-4 Sampling distribution of the sample mean difference.

	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



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 $\hat{\Theta}$.

Mean Squared Error

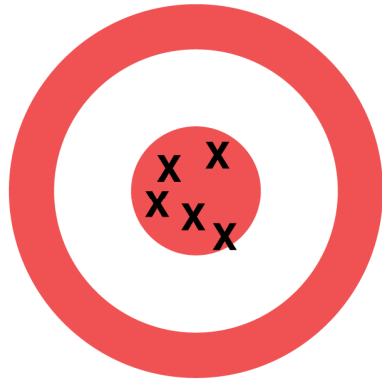
The **mean squared error** of an estimator $\hat{\theta}$ of the parameter θ is defined as:

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

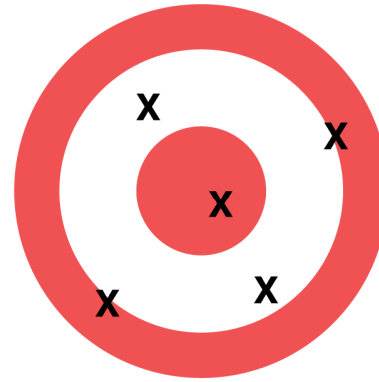
Can be rewritten as

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

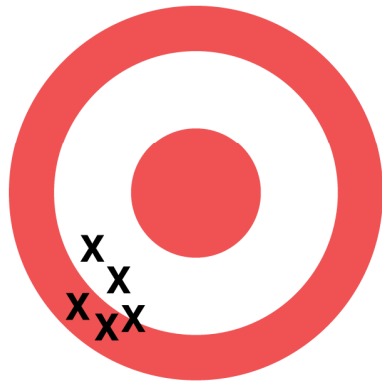
Bias vs Noise



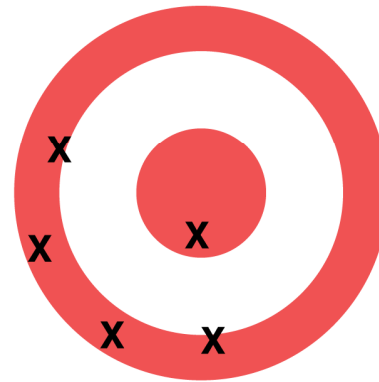
ACCURATE



NOISY



BIASED



BIASED & NOISY

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

$$E(s^2) = E\left(\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}\right) = \frac{1}{n-1} E\left[\sum_{i=1}^n (X_i^2 + \bar{X}^2 - 2\bar{X}X_i)\right] =$$
$$= \frac{1}{n-1} E\left[\sum_{i=1}^n X_i^2 + n\bar{X}^2 - 2\bar{X}n\bar{X}\right] =$$

$$= \frac{1}{n-1} E\left(\sum_{i=1}^n X_i^2 - n\bar{X}^2\right) = \frac{1}{n-1} (nE(X_i^2) - nE(\bar{X}^2))$$

$$= \frac{1}{n-1} \left(n(\mu^2 + \sigma^2) - n\left(\mu^2 + \frac{\sigma^2}{n}\right) \right) = \frac{n-1}{n-1} \sigma^2 = \underline{\underline{\sigma^2}}$$

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

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

3000

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: $\mu = \int_{-\infty}^{+\infty} xf(x) dx$
 - 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

