Probability and Statistics: Lecture-16

Monsoon-2020

by Pawan Kumar (IIIT, Hyderabad)
on September 16, 2020

» Table of contents

» Example of Poisson Distribution...

» Example of Poisson Distribution...

Example

We want to send a bit string of length $n = 10^4$.

» Example of Poisson Distribution...

Example

We want to send a bit string of length $n = 10^4$. The probability that a bit can be corrupted is $p = 10^{-6}$.

» Example of Poisson Distribution...

$n >> 1$ Binomial (n, p)

Example

We want to send a bit string of length $n = 10^4$. The probability that a bit can be corrupted is $p = 10^{-6}$. What is the probability that the message will arrive uncorrupted?

$p << 1$

$\begin{cases} n \to \infty \\ p \to \text{small} \end{cases}$

» Example of Poisson Distribution...

Example

We want to send a bit string of length $n = 10^4$. The probability that a bit can be corrupted is $p = 10^{-6}$. What is the probability that the message will arrive uncorrupted?

* We can solve this usung a Poisson with $\lambda = np = 10^4 10^{-6}$

» Example of Poisson Distribution...

Example

We want to send a bit string of length $n = 10^4$. The probability that a bit can be corrupted is $p = 10^{-6}$. What is the probability that the message will arrive uncorrupted?

* We can solve this usung a Poisson with $\lambda = np = 10^4 10^{-6}$
* With $X \sim \text{Poisson}(0.01)$ be the number of corrupted bits

» Example of Poisson Distribution...

Example

We want to send a bit string of length $n = 10^4$. The probability that a bit can be corrupted is $p = 10^{-6}$. What is the probability that the message will arrive uncorrupted?

* We can solve this usung a Poisson with $\lambda = np = 10^4 10^{-6}$
* With $X \sim \text{Poisson}(0.01)$ be the number of corrupted bits
* Using the PMF for Poisson

» Example of Poisson Distribution...

Example

We want to send a bit string of length $n = 10^4$. The probability that a bit can be corrupted is $p = 10^{-6}$. What is the probability that the message will arrive uncorrupted?

- We can solve this usung a Poisson with $\lambda = np = 10^4 10^{-6}$
- With $X \sim \text{Poisson}(0.01)$ be the number of corrupted bits
- Using the PMF for Poisson

no corrupted bits

$$
\begin{aligned}
P(X = 0) &= \frac{\lambda^0}{0!} e^{-\lambda} \\
&= \frac{0.01^0}{0!} e^{-0.01} \\
&\approx 0.9900498
\end{aligned}
$$

$x=0$

99% probab.

10^4

- We could have solved this using Binomial $\underbrace{X \sim \text{Binomial}(10^4, 10^{-6})}_{\alpha}$. This may have been harder to compute!

» Example of Poisson Distribution...

Example

We want to send a bit string of length $n = 10^4$. The probability that a bit can be corrupted is $p = 10^{-6}$. What is the probability that the message will arrive uncorrupted?

* We can solve this usung a Poisson with $\lambda = np = 10^4 10^{-6}$
* With $X \sim \text{Poisson}(0.01)$ be the number of corrupted bits
* Using the PMF for Poisson

$$\begin{aligned} P(X = 0) &= \frac{\lambda^i}{i!} e^{-\lambda} \\ &= \frac{0.01^0}{0!} e^{-0.01} \\ &\approx 0.9900498 \end{aligned}$$

* We could have solved this using Binomial $X \sim \text{Binomial}(10^4, 10^{-6})$. This may have been harder to compute!
* When n large, and p small: can use Poisson!

» Another Example of Poisson Distribution...

» Another Example of Poisson Distribution...

Example of Poisson Distribution

The Poisson distribution is often used to model the number of events that occur independently *at any* time in an interval of time or space, with *a* constant average rate.

» Another Example of Poisson Distribution...

Example of Poisson Distribution

The Poisson distribution is often used to model the number of events that occur independently at any time in an interval of time or space, with a constant average rate. Earthquakes are a good example of this.

» Another Example of Poisson Distribution...

Example of Poisson Distribution

The Poisson distribution is often used to model the number of events that occur independently at any time in an interval of time or space, with a constant average rate. Earthquakes are a good example of this. Suppose there are an average of 2.8 major earthquakes in the world each year.

» Another Example of Poisson Distribution...

Example of Poisson Distribution

The Poisson distribution is often used to model the number of events that occur independently at any time in an interval of time or space, with a constant average rate. Earthquakes are a good example of this. Suppose there are an average of 2.8 major earthquakes in the world each year. What is the probability of getting more than one major earthquake next year?

$X \sim \text{Poisson}(2.8)$, $X = \#$ earthquake

$P(\boxed{X > 1})$

$= 1 - P(X=0) - P(X=1)$

$R_X = \{0, 1, 2, \ldots\}$

» Another Example of Poisson Distribution...

Example of Poisson Distribution

The Poisson distribution is often used to model the number of events that occur independently at any time in an interval of time or space, with a constant average rate. Earthquakes are a good example of this. Suppose there are an average of 2.8 major earthquakes in the world each year. What is the probability of getting more than one major earthquake next year?

* Let $X \sim \text{Poisson}(2.8)$ be the number of major earthquakes next year. We want $P(X > 1)$.

» Another Example of Poisson Distribution...

Example of Poisson Distribution

The Poisson distribution is often used to model the number of events that occur independently at any time in an interval of time or space, with a constant average rate. Earthquakes are a good example of this. Suppose there are an average of 2.8 major earthquakes in the world each year. What is the probability of getting more than one major earthquake next year?

- Let $X \sim \text{Poisson}(2.8)$ be the number of major earthquakes next year. We want $P(X > 1)$.
- We can use

$$P(X > 1) = 1 - P(X = 0) - P(X = 1).$$

Using PMF for Poisson

» Another Example of Poisson Distribution...

Example of Poisson Distribution

The Poisson distribution is often used to model the number of events that occur independently at any time in an interval of time or space, with a constant average rate. Earthquakes are a good example of this. Suppose there are an average of 2.8 major earthquakes in the world each year. What is the probability of getting more than one major earthquake next year?

* Let $X \sim \text{Poisson}(2.8)$ be the number of major earthquakes next year. We want $P(X > 1)$.
* We can use

$$P(X > 1) = 1 - P(X = 0) - P(X = 1).$$

Using PMF for Poisson

$$\begin{aligned} P(X > 1) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - e^{-2.8}\frac{2.8^0}{0!} - e^{-2.8}\frac{2.8^1}{1!} \\ &= 1 - e^{-2.8} - 2.8e^{-2.8} \\ &\approx 1 - 0.06 - 0.17 = 0.77 \end{aligned}$$

77%

» Expectation of Poisson Distribution...

Expectation of Poisson

Let $X \sim Poisson(\lambda)$, with PMF given by

$E[X] = \sum x\, p(x)$

$$P_X(k) = \begin{cases} \frac{e^{-\lambda}\lambda^k}{k!}, & k \in R_X \\ 0 & \text{otherwise} \end{cases}$$

Show that the expectation $E[X] = \lambda$.

» Expectation of Poisson Distribution...

Recall Calculus / Real Analysis

Expectation of Poisson

Let $X \sim Poisson(\lambda)$, with PMF given by

$$P_X(k) = \begin{cases} \frac{e^{-\lambda}\lambda^k}{k!}, & k \in R_X \\ 0 & \text{otherwise} \end{cases}$$

Show that the expectation $E[X] = \lambda$.

If $f: \mathbb{R} \to \mathbb{R}$ is infinitely diff., then f can be expanded around $x = a$

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \dots$$

Proof

$$E[X] = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!}$$

$$= e^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^{z+1}}{z!} = \lambda e^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} = \lambda.$$

$x - 1 = z$

$x = 0$

$e^x : e^0 + \frac{e^0}{1} x + \frac{e^0 x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$

e^{λ}

Taylor's expansion

» Expectation of Binomial Distribution...

» Expectation of Binomial Distribution...

Expectation of a Binomial Distribution

Show that $E[X] = np$, where $X \sim \text{Binomial}(n, p), 0 < p < 1$ with PMF given as follows

$$P_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{for } k = 0, 1, 2, \cdots, n \\ 0 & \text{otherwise,} \end{cases}$$

Expectation of a Binomial Distribution

Show that $E[X] = np$, where $X \sim \text{Binomial}(n,p), 0 < p < 1$ with PMF given as follows

$$P_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{for } k = 0, 1, 2, \cdots, n \\ 0 & \text{otherwise,} \end{cases}$$

Proof

0

$$E[X] = \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x} = \sum_{x=1}^{n} x \binom{n}{x} p^x q^{n-x}.$$

P_X PMF

For $0 < x \leq n$, $$x\binom{n}{x} = x\frac{n!}{(n-x)!x!} = \frac{n!}{(n-x)!(x-1)!} = n\binom{n-1}{x-1},$$

$$\Longrightarrow \quad E[X] = \sum_{x=1}^{n} n\binom{n-1}{x-1} p^x q^{n-x} = \sum_{z=0}^{n-1} n\binom{n-1}{z} p^{z+1} q^{n-1-z} = np.$$

$x - 1 = z$

take n & p common

» Variance

Motivation for Variance

* The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:

Motivation for Variance

- The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  - You get Rs 1000

» Variance

Motivation for Variance

- The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  - You get Rs 1000
  - A fair coin is tossed, and if it is head you get Rs 2000

» Variance

Motivation for Variance

- The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  - You get Rs 1000
  - A fair coin is tossed, and if it is head you get Rs 2000
  - A number is chosen from 1 to 1000 and if you can guess this, you get Rs 1 Million

» Variance

Motivation for Variance

* The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  * You get Rs 1000
  * A fair coin is tossed, and if it is head you get Rs 2000
  * A number is chosen from 1 to 1000 and if you can guess this, you get Rs 1 Million
* Let X be the discrete random variable corresponding to your winnings. In all three cases, $E(X) = Rs\ 1000$, but you will agree that your chances of actually getting any money are quite different in all three cases.

» Variance

Motivation for Variance

* The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  * You get Rs 1000
  * A fair coin is tossed, and if it is head you get Rs 2000
  * A number is chosen from 1 to 1000 and if you can guess this, you get Rs 1 Million
* Let X be the discrete random variable corresponding to your winnings. In all three cases, $E(X) = Rs\ 1000$, but you will agree that your chances of actually getting any money are quite different in all three cases.
* The three cases differ in the spread of PMF. This can be measured by Variance.

» Variance and Standard Deviation...

» Variance and Standard Deviation...

Variance

Let $\mu = E[X]$. The variance $\mathrm{Var}(X)$ of X is defined by

$E[g(x)] = \sum_x g(x) p_x(x)$

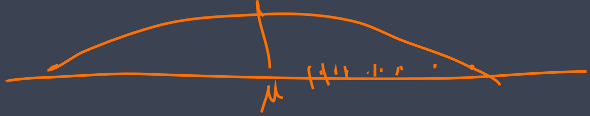
Variance

Let $\mu = E[X]$. The variance Var(X) of X is defined by

sometimes use p_x or f_X

$$\mathrm{Var}(X) = E(\underbrace{(X-\mu)^2}_{g(x)}) = \sum_x (x-\mu)^2 f_X(x)$$

» Variance and Standard Deviation...

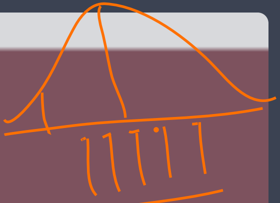


Variance

Let $\mu = E[X]$. The variance $\mathrm{Var}(X)$ of X is defined by

$$\mathrm{Var}(X) = E((X-\mu)^2) = \sum_x (x-\mu)^2 f_X(x)$$

0



The positive square root is called the standard deviation and it is usually denoted by σ, hence

Variance

Let $\mu = E[X]$. The variance $\mathrm{Var}(X)$ of X is defined by

$$\mathrm{Var}(X) = E((X-\mu)^2) = \sum_x (x-\mu)^2 f_X(x)$$

The positive square root is called the standard deviation and it is usually denoted by σ, hence

$$\sigma(X) = \sqrt{\sum_x (x-\mu)^2 f_X(x)} = \sqrt{\mathrm{Var}(X)}.$$

» Example of Computing Variance...

» Example of Computing Variance...

$R_X = \{1, 2, 3, 4, 5, 6\}$

x

Example

Let X be the value on one roll of a 6-sided die. Recall that $\underbrace{E[X]}_{\mu} = 7/2$. What is $Var(X)$?

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

$$= \sum_x (x - 7/2)^2 \, p_X(x)$$

$$= \sum_x (x-7/2)^2 \cdot \frac{1}{6}$$

$$= (1-7/2)^2 \cdot \frac{1}{6} + (2-7/2)^2 \cdot \frac{1}{6} + \cdots + (6-7/2)^2 \cdot \frac{1}{6}$$

$$=$$

Motivation for Variance

* The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:

» Variance

Motivation for Variance

- The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  - You get Rs 1000

» Variance

Motivation for Variance

* The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  * You get Rs 1000
  * A fair coin is tossed, and if it is head you get Rs 2000

» Variance

Motivation for Variance

* The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  * You get Rs 1000
  * A fair coin is tossed, and if it is head you get Rs 2000
  * A number is chosen from 1 to 1000 and if you can guess this, you get Rs 1 Million

» Variance

Motivation for Variance

* The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  * You get Rs 1000
  * A fair coin is tossed, and if it is head you get Rs 2000
  * A number is chosen from 1 to 1000 and if you can guess this, you get Rs 1 Million

» Variance

Motivation for Variance

* The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  * You get Rs 1000
  * A fair coin is tossed, and if it is head you get Rs 2000
  * A number is chosen from 1 to 1000 and if you can guess this, you get Rs 1 Million

Calculate Variance

* **First case:** There is only one outcome and it is the mean, variance is 0

» Variance

1/2 1/2

$R_X = \{0, 2000\}$

Motivation for Variance

* The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  * You get Rs 1000
  * A fair coin is tossed, and if it is head you get Rs 2000
  * A number is chosen from 1 to 1000 and if you can guess this, you get Rs 1 Million

Calculate Variance

* **First case:** There is only one outcome and it is the mean, variance is 0
* **Second case:** $Var(X) = \frac{1}{2}(2000 - 1000)^2 + \frac{1}{2}(0 - 1000)^2 = 10^6$

$x - \mu$ $\quad$ $x - \mu$

0

» Variance

Motivation for Variance

- The expectation value $E(X)$ (also called the mean) of a discrete random variable is a rather coarse measure of how X is distributed. For example, consider the following three situations:
  - You get Rs 1000
  - A fair coin is tossed, and if it is head you get Rs 2000
  - A number is chosen from 1 to 1000 and if you can guess this, you get Rs 1 Million

Calculate Variance

- **First case:** There is only one outcome and it is the mean, variance is 0
- **Second case:** $\mathrm{Var}(X) = \frac{1}{2}(2000 - 1000)^2 + \frac{1}{2}(0 - 1000)^2 = 10^6$
- **Third case:** $\mathrm{Var}(X) = 10^{-3}(10^6 - 10^3)^2 + 999 \times 10^{-3}(0 - 10^3)^2 \approx 10^9$

» Remarks on Expectation, Variance, Standard Deviation...

» Remarks on Expectation, Variance, Standard Deviation...

Example

If you are told that the average starting salary for someone working at Company Statistix is Rs 70,000, you may think, "Wow! That's great."

» Remarks on Expectation, Variance, Standard Deviation...

Example

If you are told that the average starting salary for someone working at Company Statistix is Rs 70,000, you may think, "Wow! That's great." But if the standard deviation for starting salaries at Company Statistix is Rs 20,000, that's a lot of variation in terms of how much money you can make,

» Remarks on Expectation, Variance, Standard Deviation...

Example

If you are told that the average starting salary for someone working at Company Statistix is Rs 70,000, you may think, "Wow! That's great." But if the standard deviation for starting salaries at Company Statistix is Rs 20,000, that's a lot of variation in terms of how much money you can make, so the average starting salary of Rs 70,000 isn't as informative in the end, is it?

Example

If you are told that the average starting salary for someone working at Company Statistix is Rs 70,000, you may think, "Wow! That's great." But if the standard deviation for starting salaries at Company Statistix is Rs 20,000, that's a lot of variation in terms of how much money you can make, so the average starting salary of Rs 70,000 isn't as informative in the end, is it?

* If the standard deviation was only Rs 5,000, you would have a much better idea of what to expect for a starting salary at that company

» Remarks on Expectation, Variance, Standard Deviation...

Example

If you are told that the average starting salary for someone working at Company Statistix is Rs 70,000, you may think, "Wow! That's great." But if the standard deviation for starting salaries at Company Statistix is Rs 20,000, that's a lot of variation in terms of how much money you can make, so the average starting salary of Rs 70,000 isn't as informative in the end, is it?

* If the standard deviation was only Rs 5,000, you would have a much better idea of what to expect for a starting salary at that company
* Without calculating standard deviation, you can't get a handle on whether the data are close to the average

Example

If you are told that the average starting salary for someone working at Company Statistix is Rs 70,000, you may think, "Wow! That's great." But if the standard deviation for starting salaries at Company Statistix is Rs 20,000, that's a lot of variation in terms of how much money you can make, so the average starting salary of Rs 70,000 isn't as informative in the end, is it?

* If the standard deviation was only Rs 5,000, you would have a much better idea of what to expect for a starting salary at that company
* Without calculating standard deviation, you can't get a handle on whether the data are close to the average

Same ~~Standard Deviation~~ [average] Does Not Imply Same ~~Dataset~~ [Std dev]

For example, the data sets 199, 200, 201 and 0, 200, 400 both have the same average (200) yet they have very different standard deviations.

» Remarks on Expectation, Variance, Standard Deviation...

Example

If you are told that the average starting salary for someone working at Company Statistix is Rs 70,000, you may think, “Wow! That’s great.” But if the standard deviation for starting salaries at Company Statistix is Rs 20,000, that’s a lot of variation in terms of how much money you can make, so the average starting salary of Rs 70,000 isn’t as informative in the end, is it?

* If the standard deviation was only Rs 5,000, you would have a much better idea of what to expect for a starting salary at that company
* Without calculating standard deviation, you can’t get a handle on whether the data are close to the average

Same Standard Deviation Does Not Imply Same Dataset

For example, the data sets 199, 200, 201 and 0, 200, 400 both have the same average (200) yet they have very different standard deviations. The first data set has a very small standard deviation (s=1) compared to the second data set (s=200).

» Another expression for the variance

Theorem (Another Expression for Variance)

If X is a discrete random variable with mean μ, then

$$\mathrm{Var}(X) = E[X^2] - \mu^2$$

» Another expression for the variance

Theorem (Another Expression for Variance)

If X is a discrete random variable with mean μ, then

$$\mathrm{Var}(X) = E[X^2] - \mu^2$$

Proof

$$\mathrm{Var}(X) = \sum_x (x-\mu)^2 p_X(x) = \sum_x (x^2 - 2\mu x + \mu^2) p_X(x)$$

$$= \sum_x x^2 p_X(x) - 2\mu \sum_x x p_X(x) + \mu^2 \sum_x p_X(x)$$

1

$$= E[X^2] - 2\mu E[X] + \mu^2 = E[X^2] - \mu^2$$

μ

» Properties of Variance...

Theorem

Let X be a discrete random variable and α a constant. Then

$$\mathsf{Var}(\alpha X) = \alpha^2 \mathsf{Var}(X) \quad \text{and} \quad \mathsf{Var}(X + \alpha) = \mathsf{Var}(X)$$

$\boxed{E[X] = \mu}$

$E[\alpha X] = \alpha\mu$

$\mathrm{Var}(\alpha X) = E[(\alpha X)^2] - \alpha^2\mu^2$

$= \alpha^2 E(X^2) - \alpha^2\mu^2$

$\cancel{\mathrm{Var}(\alpha X) = E[(\alpha X - \alpha\mu)^2]}$

$= \alpha^2[E(X^2) - \mu^2]$

$\alpha^2\,\mathrm{Var}(X)$

$E[X+\alpha] = E[X] + E[\alpha]$

$= \mu + \alpha$

$\mathrm{Var}(X+\alpha) = E[(X+\alpha)^2] - (\mu+\alpha)^2$

$= \mathrm{Var}(X)$

» Computing Variance: Binomial

$\text{Binomial}(n, p)$ $\qquad E[X] = np$

Variance of Binomial Distribution

Let $X \sim \text{Binomial}(n, p)$. Then the variance $\text{Var}(X) = np(1-p)$.

$$\text{Var}(X) = E(X^2) - \mu^2 = \underbrace{E[X(X-1)]} + \underbrace{E[X]} - \underbrace{E[X]}^2$$

$$E[X(X-1)] = \sum_{x=0}^{n} \underbrace{x(x-1)} \binom{n}{x} p^x (1-p)^{n-x}$$

$$\sum_{x=2}^{n} \frac{n!}{\underbrace{(x-2)!}\,(n-x)!} p^x (1-p)^{n-x}$$

» Variance of Binomial Distribution...

$$= \sum_{x=2}^{n} \frac{n!}{(x-2)!\,(n-x)!} p^{x} (1-p)^{n-x}$$

$$= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!\,(n-x)!} p^{x-2} (1-p)^{n-x}$$

$$\begin{cases} m = n-2 \\ y = x-2 \end{cases}$$

$$= n(n-1)p^2 \underbrace{\sum \frac{m!}{y!\,(m-y)!} p^{y} (1-p)^{m-y}}$$

$$= \underline{n(n-1)p^2} \left(p + (1-p) \right)^{m} = n(n-1)p^2$$

$$= E\left[X(X-1)\right]$$

$$\mathrm{Var}(X) = E[\underline{X(X-1)}] + E[X] - (E[X])^2$$

$$= n(n-1)p^2 + np - (np)^2$$

$$= \underline{\underline{np(1-p)}}$$

» Higher Order Moments...

Define nth moment

The nth moment about the mean or nth central moment of a real valued random variable X is defined as follows

$$\mu_n = E[(X - E[X])^n],$$

where E is the expectation operator.

$E[X]$

$E[X-\mu]$

$E[(X-\mu)^2]$

$E[(X-\mu)^3]$

» Higher Order Moments...

Define nth moment

The nth moment about the mean or nth central moment of a real valued random variable X is defined as follows

0

$$\mu_n = E[(X - E[X])^n],$$

$E[1] = 1$

where E is the expectation operator.

* The zeroth central moment μ_0 is 1

Define *n*th moment

The *n*th moment about the mean or *n*th central moment of a real valued random variable X is defined as follows

$$\mu_n = E[(X - E[X])^n],$$

where E is the expectation operator.

$E[(X-\mu)]$
$= E[X] - \mu$
$= \mu - \mu = 0$

- The zeroth central moment μ_0 is 1
- The first central moment μ_1 is 0 (not the same as expected value!)

» Higher Order Moments...

Define nth moment

The nth moment about the mean or nth central moment of a real valued random variable X is defined as follows

$$\mu_n = E[(X - E[X])^n],$$

where E is the expectation operator.

* The zeroth central moment μ_0 is 1
* The first central moment μ_1 is 0 (not the same as expected value!)
    * First central moment about origin is indeed expectation!

» Higher Order Moments...

Define nth moment

The *n*th moment about the mean or *n*th central moment of a real valued random variable X is defined as follows

$$\mu_n = E[(X - E[X])^n],$$

$E[(x-\mu)^2] =$

where E is the expectation operator.

- The zeroth central moment μ_0 is 1
- The first central moment μ_1 is 0 (not the same as expected value!)
  - First central moment about origin is indeed expectation!
- The second central moment μ_2 is called variance, denoted by σ^2

» Higher Order Moments...

Define nth moment

The *nth moment* about the mean or *nth central moment* of a real valued random variable X is defined as follows

$$\mu_n = E[(X - E[X])^n],$$

where E is the expectation operator.

* The zeroth central moment μ_0 is 1
* The first central moment μ_1 is 0 (not the same as expected value!)
  * First central moment about origin is indeed expectation!
* The second central moment μ_2 is called variance, denoted by σ^2

Generating Moments...

Is there a quick way to generate moments?

» Moment Generating Function...

» Moment Generating Function...

$e^{tX} = g(x)$

Moment Generating Function

The moment generating function $M_X(t)$ is the expectation value

$$M_X(t) = E[e^{tX}] = \sum_x e^{tx} p_X(x)$$

$\sum p_X(x) = 1$

$g(x)$ pmf f_X

Lemma

* $M_X(0) = 1$
* $E[X] = M_X'(0)$, where ' is the derivative w.r.t. t

$t \to 0$

derivative

$\frac{d M_X(t)}{dt} = \sum x \cdot e^{tx} p_X(x)$

$\therefore \sum x\, p_X(x) = E[X]$

$= \mu$