

CS 215

Data Analysis and Interpretation

Transformation of Random Variables

Suyash P. Awate

Topics

- Transformation of random variables
 - Univariate
 - Multivariate
- Multivariate statistics
 - Multivariate Gaussian
 - Functional form, properties
 - Relationships with linear algebra
 - Singular value decomposition (SVD) of matrices
 - Eigen decomposition of matrices
- Principal component analysis (PCA)
 - Relies on covariances of multiple random variables
 - Needn't be multivariate Gaussian
 - Special properties when data is multivariate Gaussian

Topics

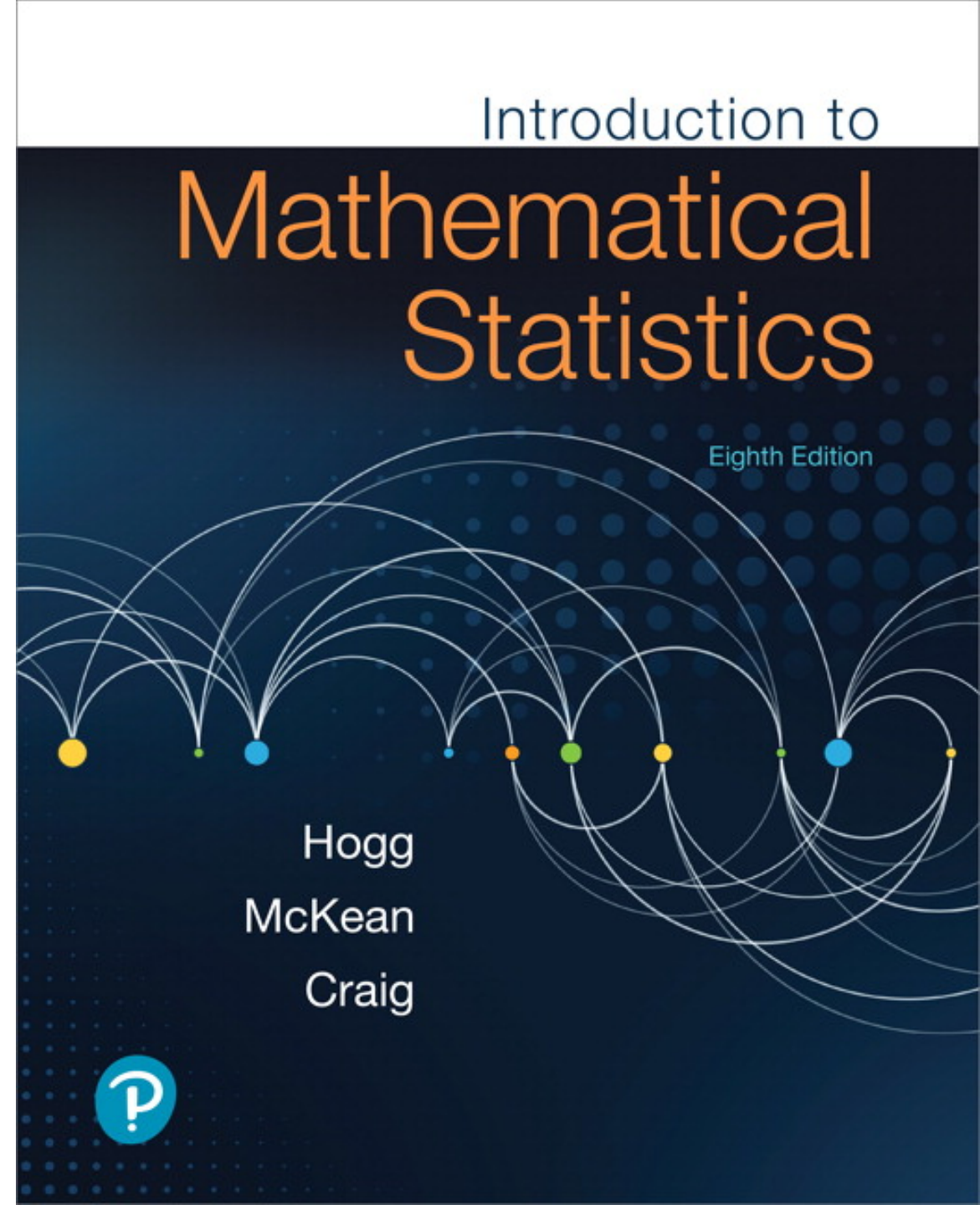
- Bayesian statistics
 - Improving over maximum-likelihood estimation
 - Prior models
 - Some concepts from information theory
- Measuring quality of estimators
 - Fisher information
 - Cramer-Rao lower bound
 - Bayesian Cramer-Rao lower bound

Course Organization

- Lectures remaining: 10
 - Last date of instruction 20 Nov
 - Holiday 30 Oct
- Tutorial sessions (2)
 - Multivariate statistics
 - Bayesian statistics
- Assignments (2; take home)
 - Multivariate statistics, PCA
 - Bayesian statistics
- Quiz (1): lecture time slot ? 10 -> 9 ?
- End-semester examination
- Online meetings: 45 minutes in regular time slots (twice each week)

Reference Book

- [Link 1](#)
- [Link 2](#)
- [Link 3](#)
- Will suggest books, tutorials, notes
- Find your favorite book



Mathematical Statistics

- Statistics (Wikipedia)
 - Discipline that concerns the collection, organization, **analysis, interpretation** and presentation of **data**
- Mathematical Statistics (Wikipedia)
 - Application of probability theory, a branch of mathematics, to statistics, as opposed to techniques for collecting statistical data
 - Specific mathematical techniques which are used for this include:
 - Mathematical analysis
 - Linear algebra
 - Stochastic analysis
 - Differential equations
 - Measure theory

Statistics

- Mark Twain
 - 1835 – 1910

A black and white portrait of Mark Twain, showing him from the chest up. He has white, wavy hair and a prominent white mustache. He is looking slightly to the right of the camera with a serious expression. The background is dark.

**“THERE ARE LIES, DAMNED LIES AND
STATISTICS.”**

MARK TWAIN

© Lifehack Quotes

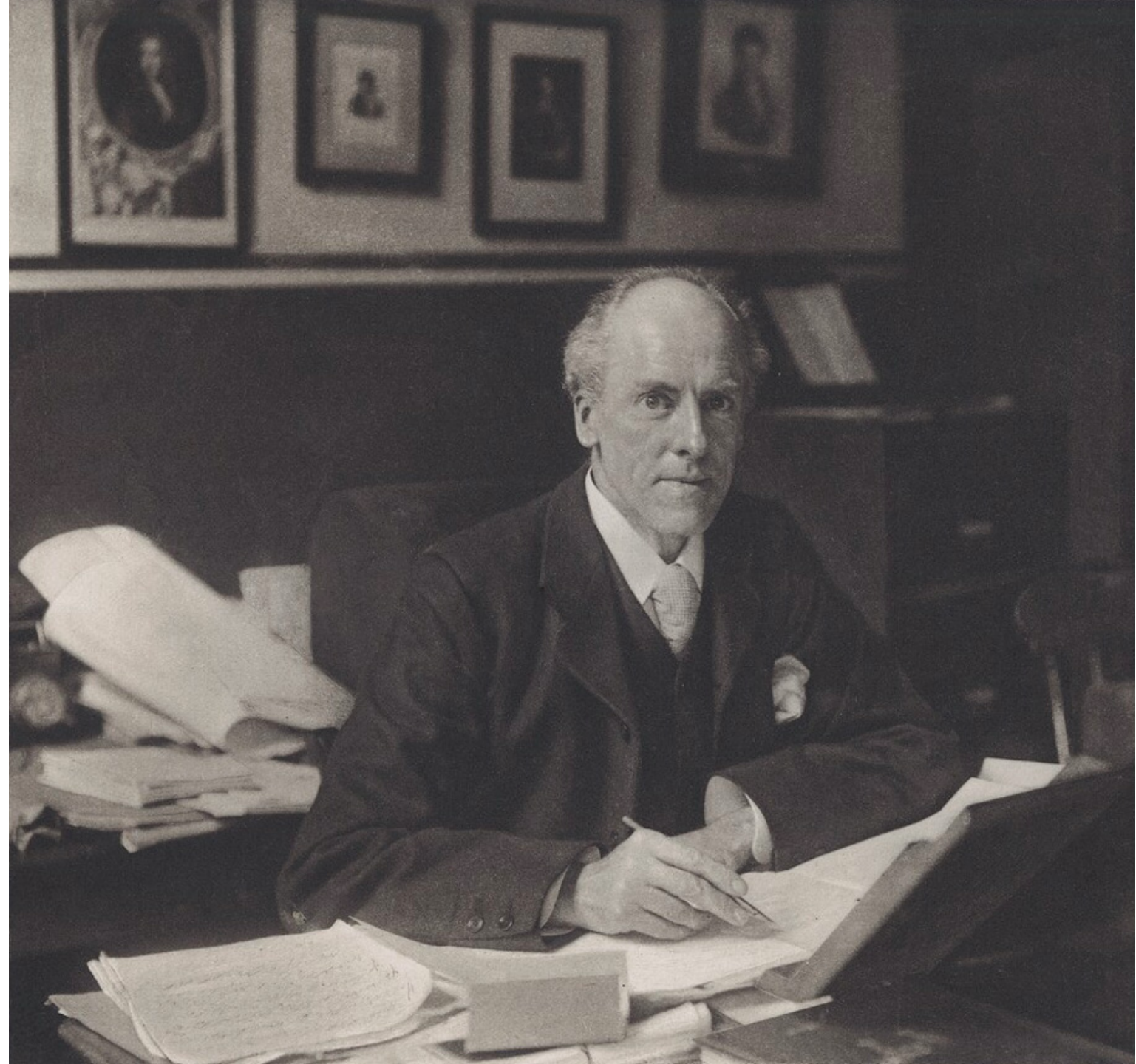
Statistics

- Ronald Fisher
 - A genius who almost single-handedly created the foundations for modern statistical science
 - 1890 – 1962



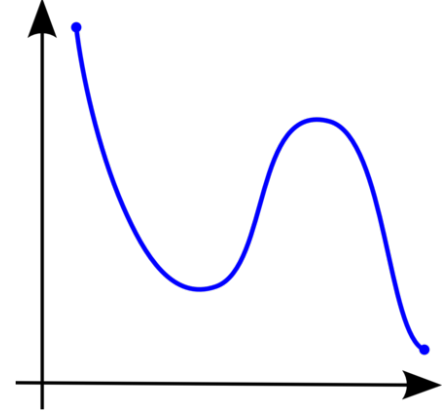
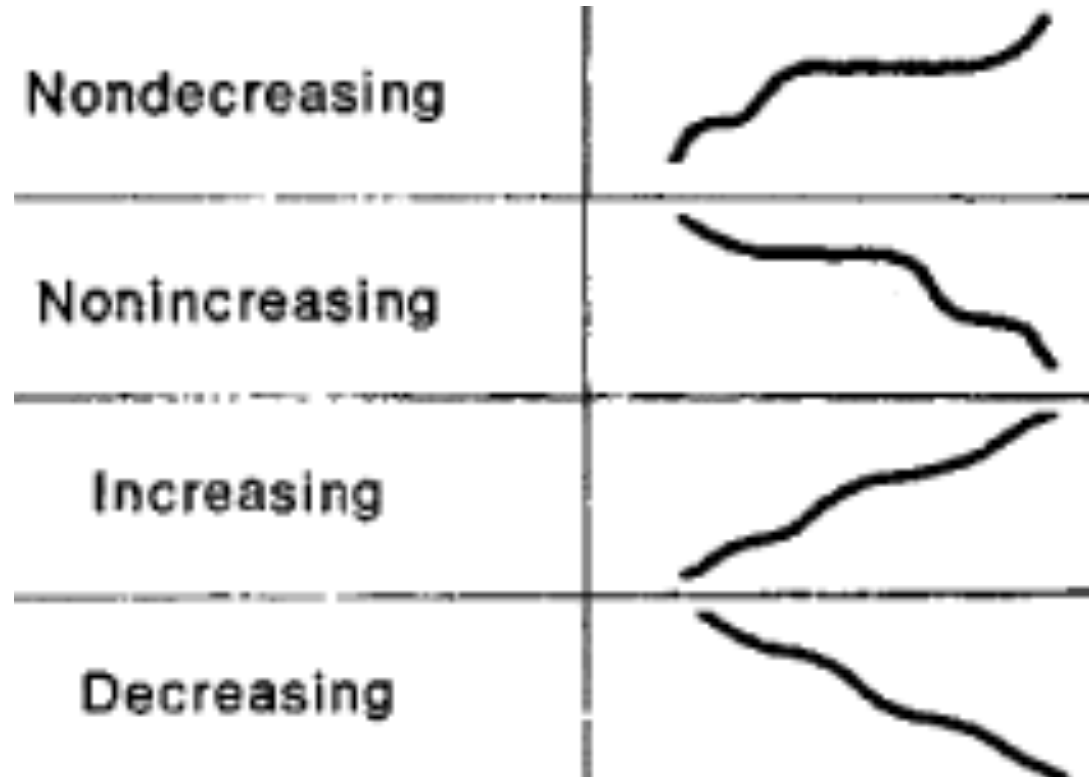
Statistics

- Karl Pearson
 - Established the discipline of mathematical statistics
 - 1857 – 1936



Transformation of Random Variables

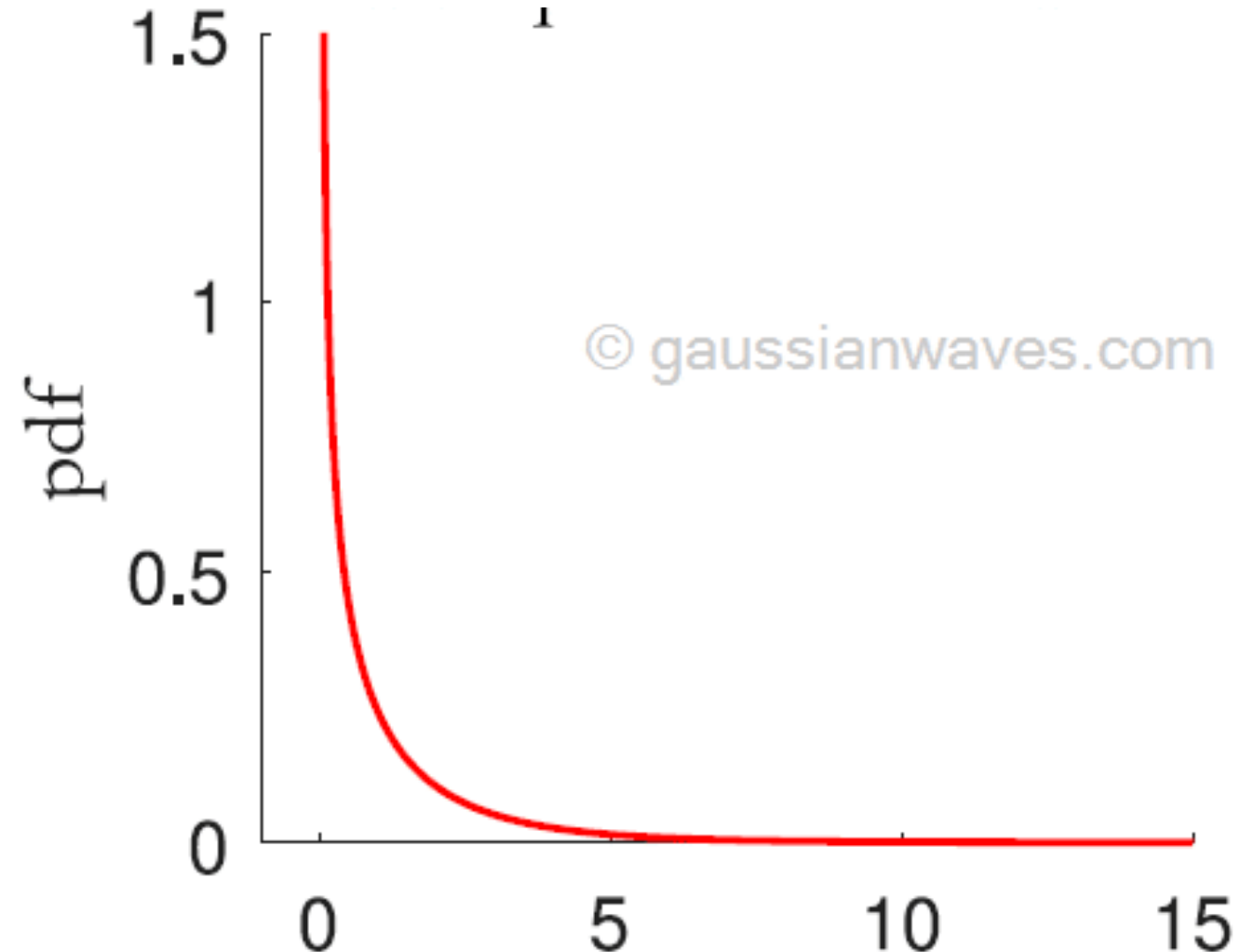
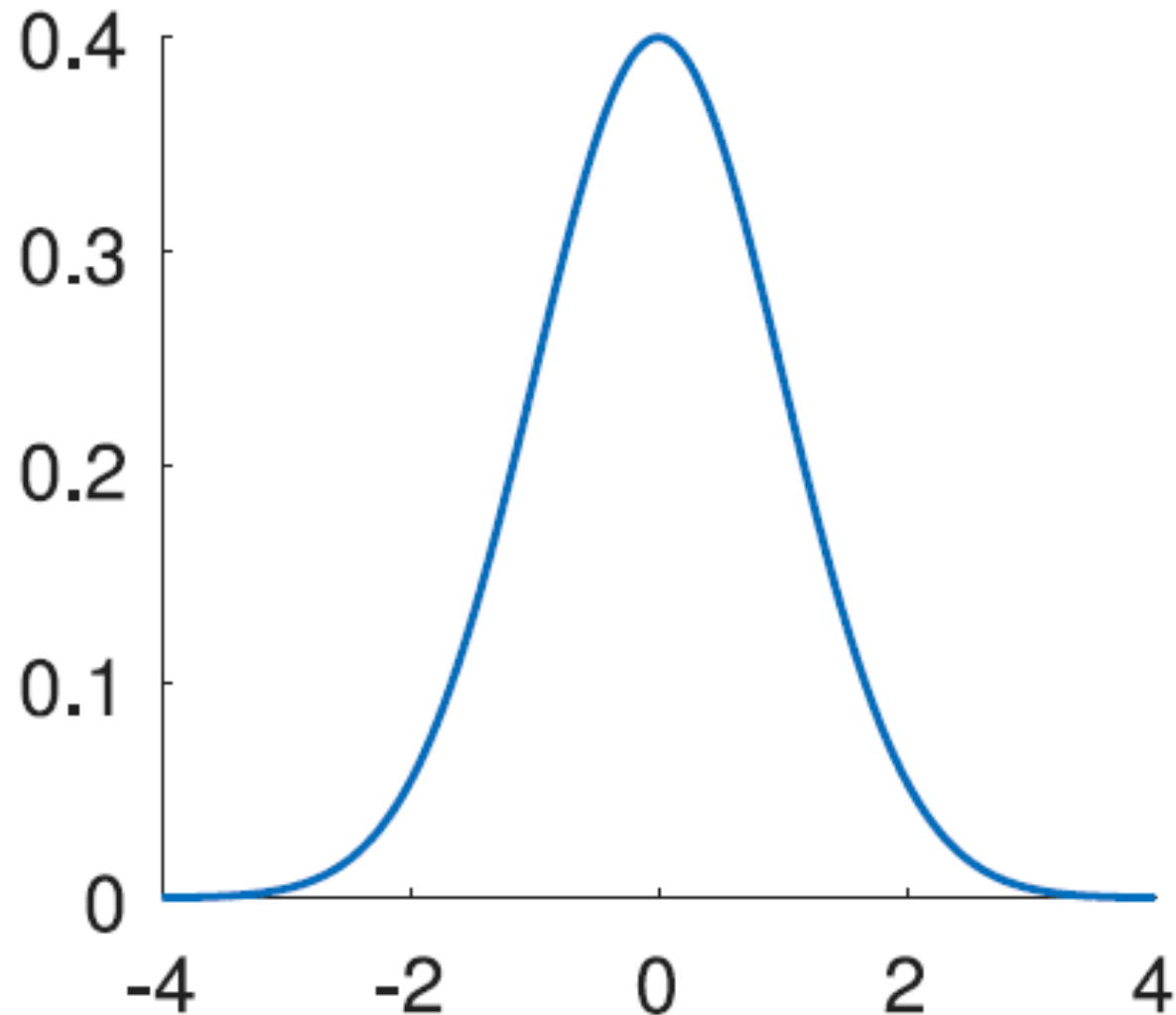
- Let X be a (continuous) random variable (RV) with probability density function (PDF) $p(X)$
- Let function $g(\cdot)$ be strictly monotonically **increasing**
 - If $a < b$, then $g(a) < g(b)$
- We will generalize/extend the class of functions later
- Consider the transformed variable $Y := g(X)$
- What is the PDF $q(Y)$ of RV Y ?



Transformation of Random Variables

- Example

- If X has a Normal PDF, then what is the PDF for $Y := X^2$?



Transformation of Random Variables

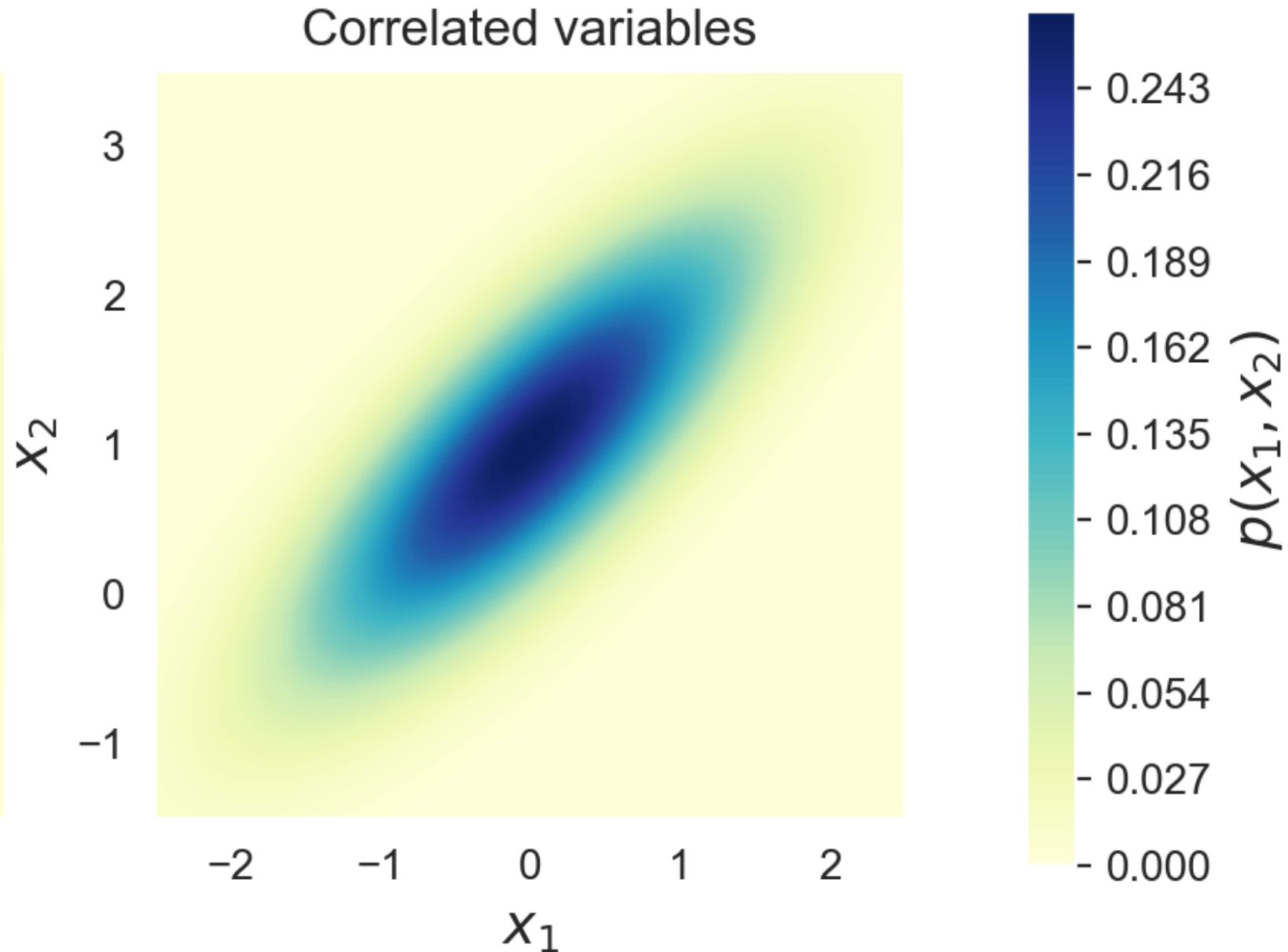
- Example

- If

- RVs U, V are independent Gaussian
 - $X_1 = aU + bV$
 - $X_2 = cU + dV$

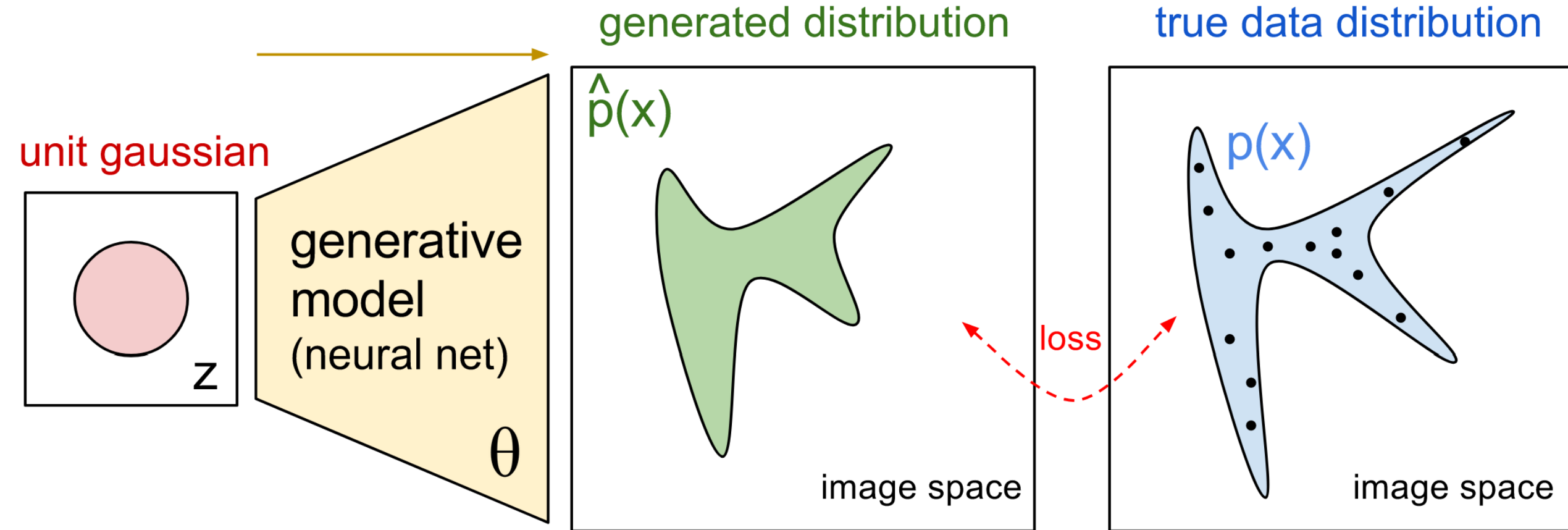
- Then

- What is $P(X_1, X_2)$?



Transformation of Random Variables

- Example



Transformation of Random Variables

- Principle of probability mass conservation
 - Consider the events $\{x : x \in (a, b)\}$ and $\{y : y \in (g(a), g(b))\}$
 - Because we assumed that $g(\cdot)$ was increasing, $P(g(a) < Y < g(b)) = P(a < X < b)$
- So, the probability “mass” of X in interval (a, b) gets mapped to the mass of Y in interval $(g(a), g(b))$

$$\text{Now, } P(g(a) < Y < g(b)) := \int_{g(a)}^{g(b)} q(y) dy$$

$$\text{Also, } P(a < X < b) := \int_a^b p(x) dx$$

- Write the second integral in terms of y , using the known relationship $y = g(x)$

Transformation of Random Variables

- We found that these probabilities $\rightarrow P(g(a) < Y < g(b)) := \int_{g(a)}^{g(b)} q(y) dy$ are equal

$$P(a < X < b) := \int_a^b p(x) dx$$

- We have, $x = g^{-1}(y)$

$$dx = \left(\frac{d}{dy} g^{-1}(y) \right) dy$$

$$\text{Then, } P(a < X < b) = \int_{g(a)}^{g(b)} p(g^{-1}(y)) \left(\frac{d}{dy} g^{-1}(y) \right) dy$$

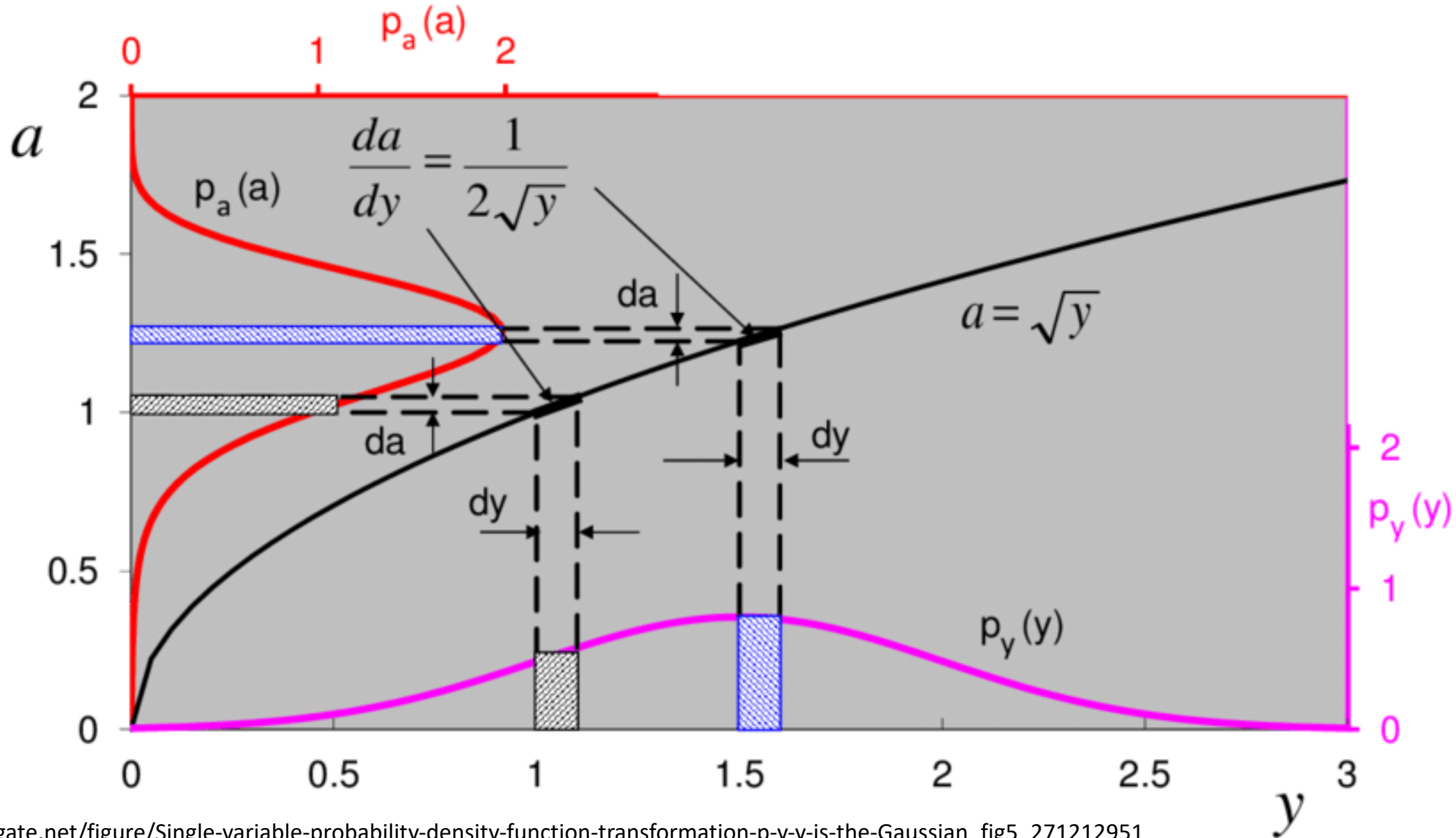
- This mass conservation holds for **every** interval (a,b), however small

$$\text{Thus, } q(y) = p(g^{-1}(y)) \frac{d}{dy} g^{-1}(y), \text{ for all } y$$

Transformation of Random Variables

- Example

- $P(Y)$
- $A := \sqrt{Y}$
- $g(\cdot)$ is $\sqrt{\cdot}$
- To find $P(A)$



Transformation of Random Variables

- We found the relationship between PDF $q(\cdot)$ of Y and PDF $p(\cdot)$ of X , in terms of the strictly-increasing transformation function $g(\cdot)$

$$q(y) = p(g^{-1}(y)) \frac{d}{dy} g^{-1}(y), \text{ for all } y$$

- If $g(\cdot)$ is strictly increasing, then:
 - $a < b \Rightarrow g(a) < g(b)$
 - Derivative of g -inverse(\cdot) is positive
 - So, the above formula holds good
- If $g(\cdot)$ is strictly decreasing, then:
 - $a < b \Rightarrow g(a) > g(b)$
 - Derivative of g -inverse(\cdot) is negative
 - What to do then ?

Transformation of Random Variables

- For convenience, to handle both cases above, we:
 - Write $q(y) = p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$
 - Taking the absolute value ensures that PDF $q(\cdot)$ is always non-negative
 - Take the integral limits to go from a smaller number to a larger number

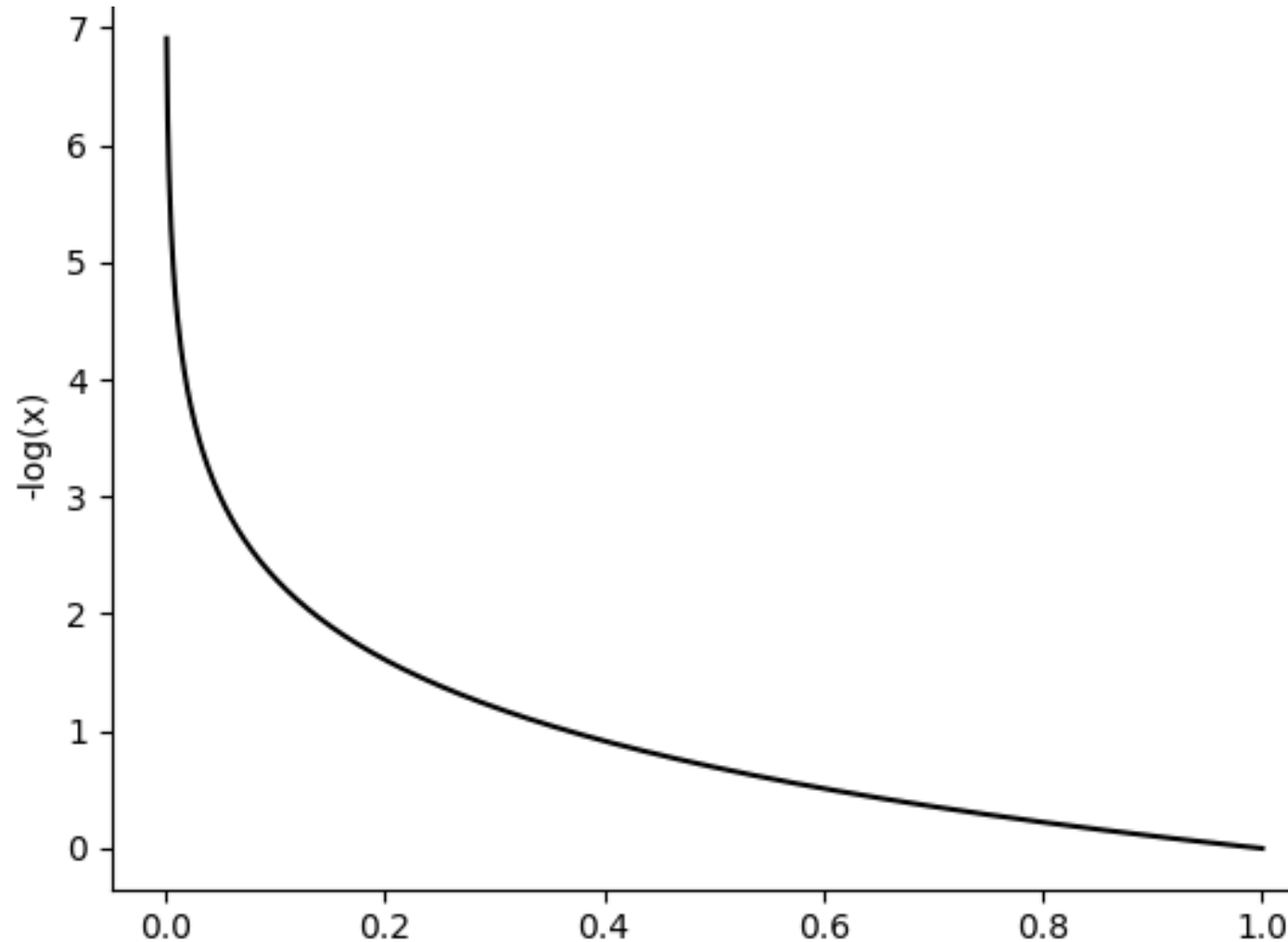
We have, $x = g^{-1}(y)$

$$dx = \left(\frac{d}{dy} g^{-1}(y) \right) dy$$

$$\text{Then, } P(a < X < b) = \int_{g(a)}^{g(b)} p(g^{-1}(y)) \left(\frac{d}{dy} g^{-1}(y) \right) dy$$

Transformation of Random Variables

- Consider a RV $X \sim U(0,1)$ (generated by the C/C++ `rand()` function)
- Consider the transformation $Y := (-1/\lambda) \log(X)$, where $\lambda > 0$
- What is $q(Y)$?



Transformation of Random Variables

- Consider a RV $X \sim U(0,1)$ (generated by the C/C++ `rand()` function)
- Consider the transformation $Y := (-1/\lambda) \log(X)$, where $\lambda > 0$
- What is $q(Y)$?

$y = -(1/\lambda) \log(x) \implies x = \exp(-\lambda y)$. This is the $g^{-1}(\cdot)$ function.

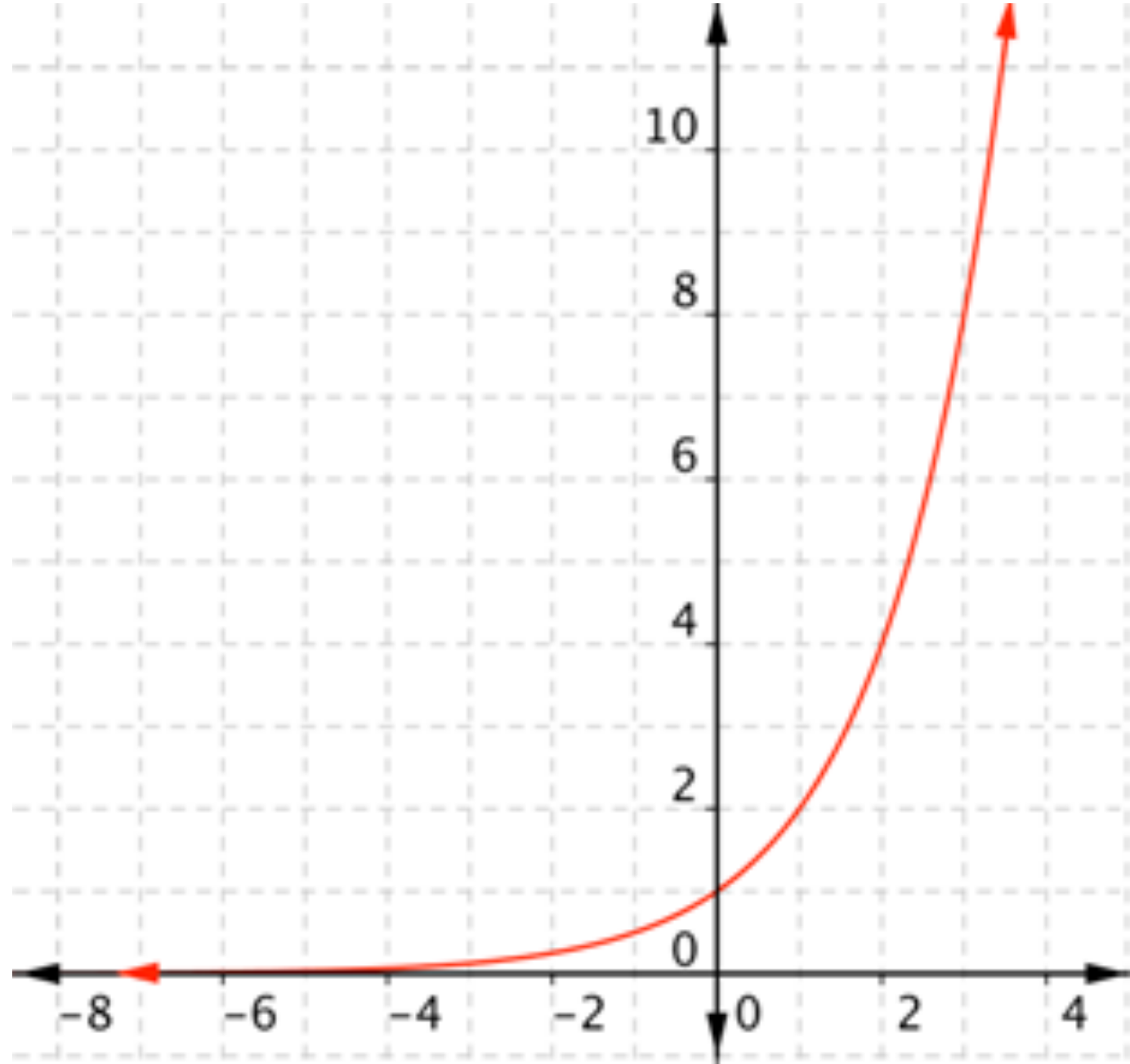
$$\left| \frac{d}{dy} g^{-1}(y) \right| = \lambda \exp(-\lambda y)$$

$$\text{So, } q(y) = p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = \lambda \exp(-\lambda y)$$

- Thus, Y has the exponential PDF with parameter λ , i.e., mean = $1/\lambda$

Transformation of Random Variables

- Consider a RV $X \sim U(-a/2, +a/2)$
- Consider $Y := \exp(X)$
- What is $q(Y)$?



Transformation of Random Variables

- Consider a RV $X \sim U(-a/2, +a/2)$
- Consider $Y := \exp(X)$
- What is $q(Y)$?

$y = \exp(x) \implies x = \log(y)$. This is the $g^{-1}(\cdot)$ function.

$$\left| \frac{d}{dy} g^{-1}(y) \right| = 1/y$$

$$\text{So, } q(y) = p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = (1/a)(1/y)$$

- Thus, Y has PDF $q(y) = 1/(ay)$ for $y \in (\exp(-a/2), \exp(a/2))$

Transformation of Random Variables

- Consider a RV $X \sim G(0,1)$ (standard Normal PDF)
- Consider $Y := aX$, with 'a' non-zero
- What is $q(Y)$?

$$y := ax \implies x = y/a \implies g^{-1}(y) = y/a$$

$$\left| \frac{d}{dy} g^{-1}(y) \right| = 1/a$$

$$q(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = p\left(\frac{y}{a}\right) \frac{1}{a} = \frac{1}{a\sqrt{2\pi}} \exp\left(-\frac{y^2}{2a^2}\right)$$

- Thus, $p(Y)$ is also a Gaussian with variance σ^2 scaled by a factor of a^2

Transformation of Random Variables

- Consider a RV $X \sim G(0, a^2)$
- Consider $Y := X + b$
- What is $q(Y)$?

$$y := b + x \implies x = y - b \implies g^{-1}(y) = y - b$$

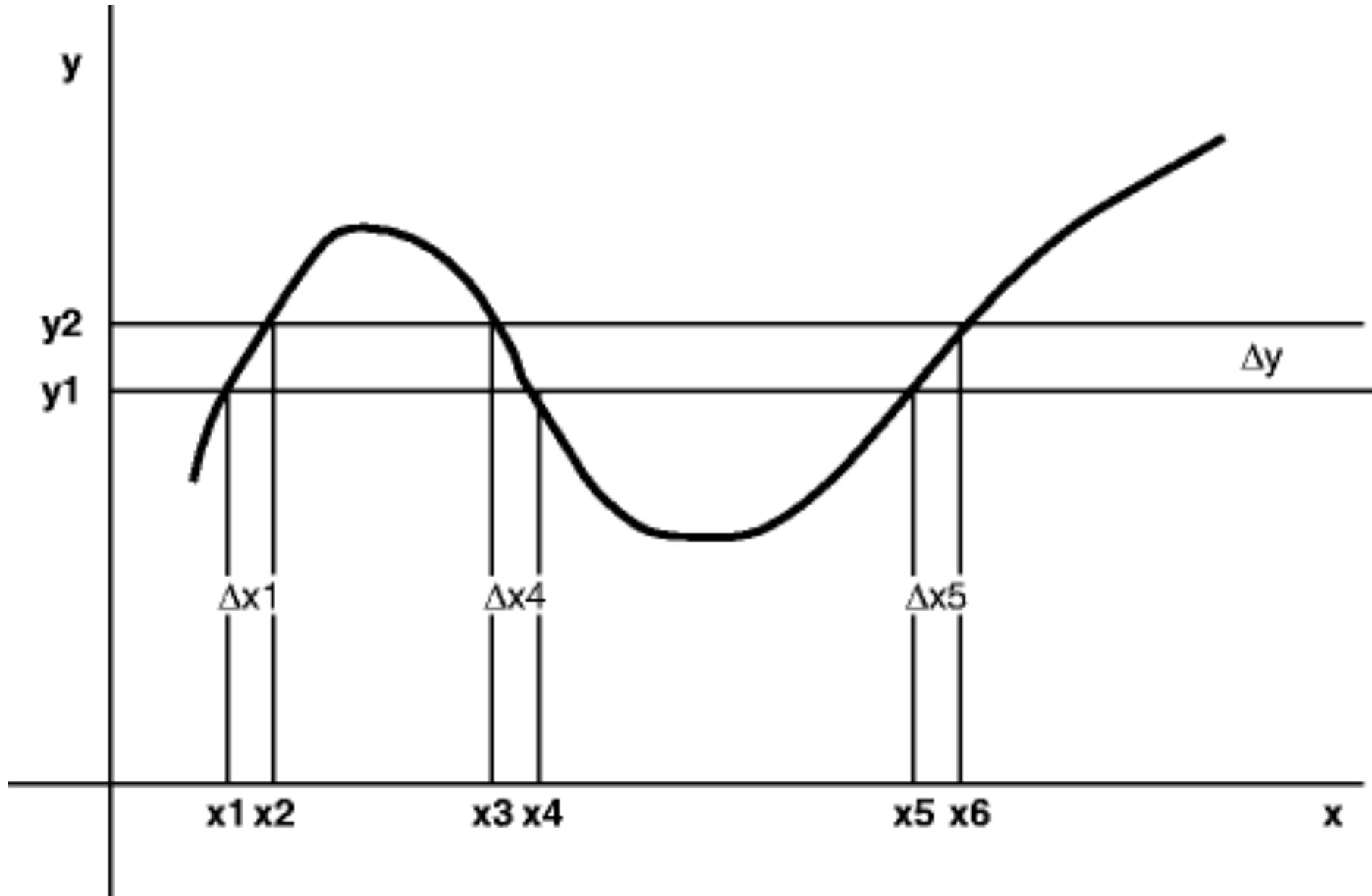
$$\left| \frac{d}{dy} g^{-1}(y) \right| = 1$$

$$q(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = p(y - b) \cdot 1 = \frac{1}{a\sqrt{2\pi}} \exp \left(-\frac{(y - b)^2}{2a^2} \right)$$

- Thus, $p(Y)$ is also a Gaussian with μ translated by b

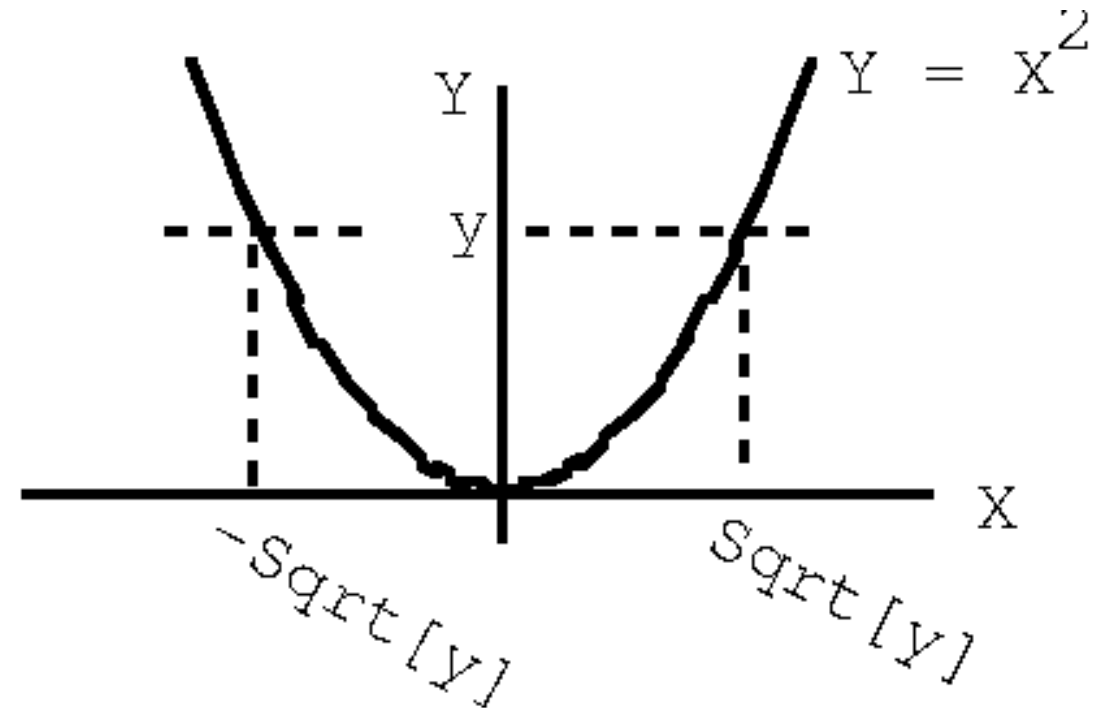
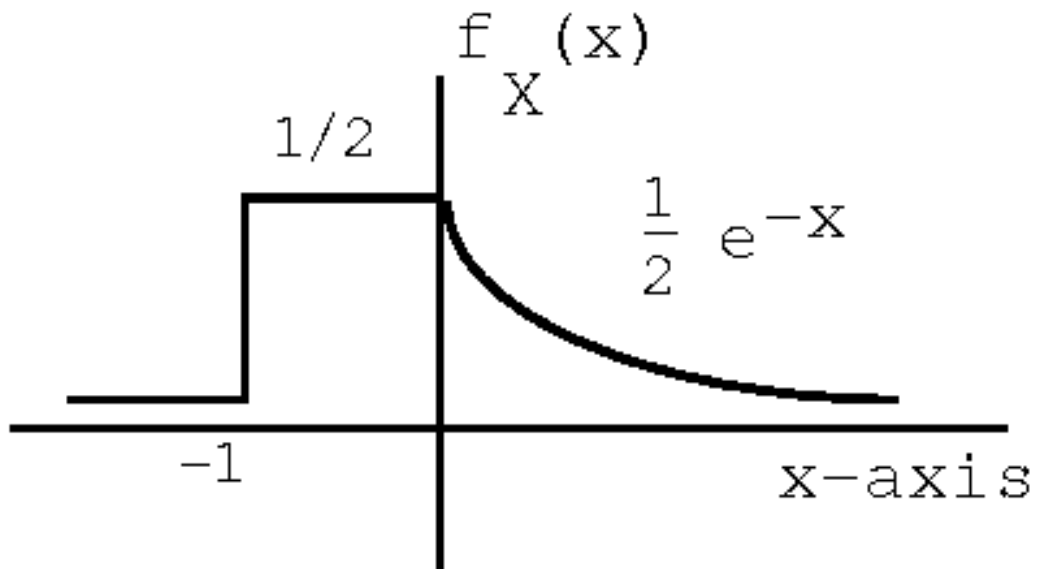
Transformation of Random Variables

- General non-monotonic functions



Transformation of Random Variables

- Consider a PDF $P(X)$ as follows:
 $P(x) := 0$ for $x \leq -1$
 $P(x) := 0.5$ for $x \in (-1, 0)$
 $P(x) := 0.5 \exp(-x)$ for $x \geq 0$
- Consider a transformation function $Y := g(X) := X^2$
- What is PDF $q(y)$ of Y ?



Transformation of Random Variables

- Consider a PDF $P(X)$ as follows:
 $P(x) := 0$ for $x \leq -1$
 $P(x) := 0.5$ for $x \in (-1, 0)$
 $P(x) := 0.5 \exp(-x)$ for $x \geq 0$
- Consider a transformation function $Y := g(X) := X^2$
- What is PDF $q(y)$ of Y ?

$$y := x^2 \implies x = \pm\sqrt{y} \implies g^{-1}(y) = \pm\sqrt{y}$$

$$\left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{2\sqrt{y}}$$

Transformation of Random Variables

- Consider a PDF $P(X)$ as follows:
$$P(x) := 0 \text{ for } x \leq -1$$
$$P(x) := 0.5 \text{ for } x \in (-1, 0)$$
$$P(x) := 0.5 \exp(-x) \text{ for } x \geq 0$$
- Consider a transformation function $Y := g(X) := X^2$
- What is PDF $q(y)$ of Y ?

Case 1: $x \in (-1, 0)$. In this case, $g(\cdot)$ is a *decreasing* function. Mass conservation applies.

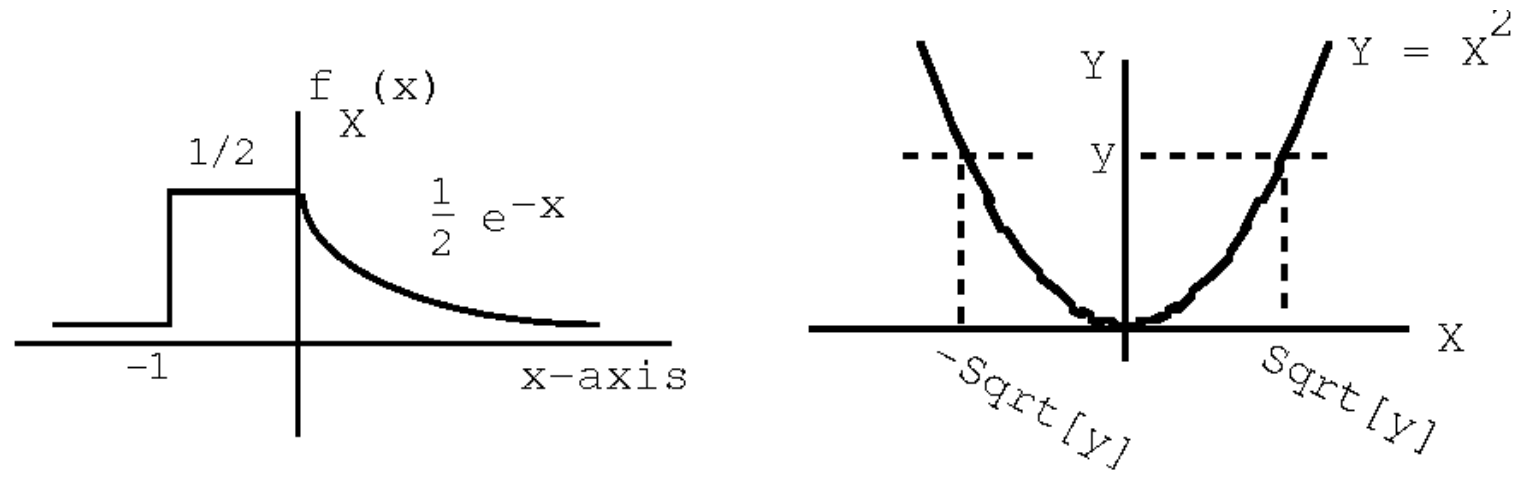
$$\text{For } y \in (0, 1) : q_1(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = (0.5) \frac{1}{2\sqrt{y}} = \frac{1}{4\sqrt{y}}$$

Case 2: $x \geq 0$. In this case, $g(\cdot)$ is a *increasing* function. Mass conservation applies.

$$\text{For } y \geq 0 : q_2(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = (0.5 \exp(-\sqrt{y})) \frac{1}{2\sqrt{y}} = \frac{\exp(-\sqrt{y})}{4\sqrt{y}}$$

Transformation of Random Variables

- Consider a PDF $P(X)$ as follows:
$$P(x) := 0 \text{ for } x \leq -1$$
$$P(x) := 0.5 \text{ for } x \in (-1, 0)$$
$$P(x) := 0.5 \exp(-x) \text{ for } x \geq 0$$
- Consider a transformation function $Y := g(X) := X^2$
- What is PDF $q(y)$ of Y ?
- Desired PDF $q(y) = q_1(y) + q_2(y)$
- In the region $y \in (0,1)$, probability mass comes from Case 1 & Case 2



Transformation of Random Variables

- Consider a PDF $P(X)$ as follows:
$$P(x) := 0 \text{ for } x \leq -1$$
$$P(x) := 0.5 \text{ for } x \in (-1, 0)$$
$$P(x) := 0.5 \exp(-x) \text{ for } x \geq 0$$
- Consider a transformation function $Y := g(X) := X^2$
- What is PDF $q(y)$ of Y ?

Thus,

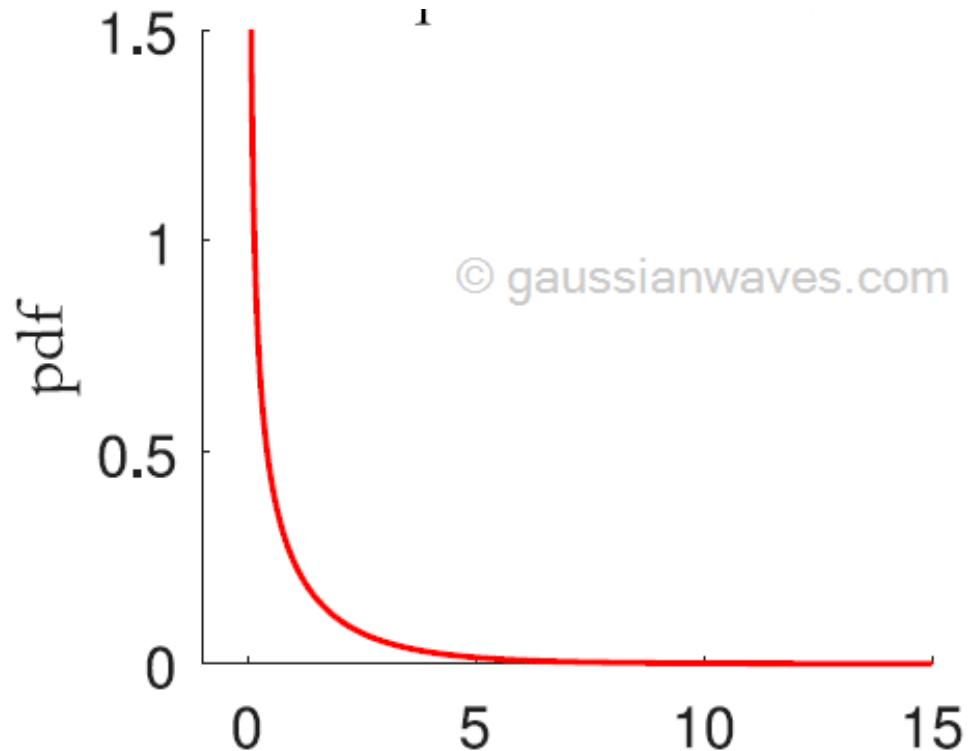
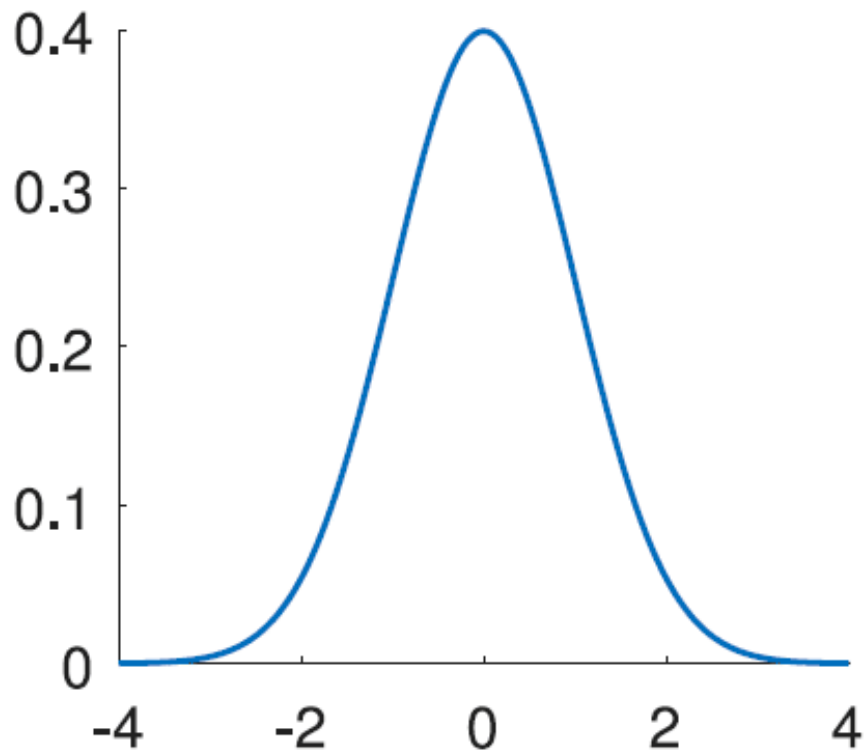
(i) for $y \in (0, 1)$, PDF $q(y) = \frac{1}{4\sqrt{y}}(1 + \exp(-\sqrt{y}))$

(ii) for $y \geq 1$, PDF $q(y) = \frac{\exp(-\sqrt{y})}{4\sqrt{y}}$

- There will be a jump discontinuity at $y = 1$,
where left limit = $(1+\exp(-1))/4$ and right limit = $\exp(-1)/4$

Transformation of Random Variables

- Let $X \sim G(0,1)$
- Let $Y := X^2$
- What is $P(Y)$, defined as the chi-square PDF ?



Transformation of Random Variables

- Let $X \sim G(0,1)$

$$y := x^2 \implies x = \pm\sqrt{y} \implies g^{-1}(y) = \pm\sqrt{y}$$

- Let $Y := X^2$

$$\left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{2\sqrt{y}}$$

- What is $P(Y)$?

- Case 1: $x \leq 0$. Here, $g(\cdot)$ is a decreasing function

$$\text{For } y \geq 0 : q_1(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = \frac{\exp(-0.5(\sqrt{y})^2)}{\sqrt{2\pi}} \frac{1}{2\sqrt{y}} = \frac{\exp(-0.5y)}{2\sqrt{y}2\pi}$$

- Case 2: $x > 0$. Here, $g(\cdot)$ is an increasing function

$$\text{For } y > 0 : q_2(y) := \frac{\exp(-0.5y)}{2\sqrt{y}2\pi}$$

- Desired chi-square PDF: $q(y) = q_1(y) + q_2(y) = (1/\sqrt{y2\pi}) \exp(-0.5y)$

Transformation of Random Variables

- Let X have a Gamma PDF,

$$P(x) = \text{Gamma}(x|\alpha, \beta) = (\beta^\alpha / \Gamma(\alpha)) x^{\alpha-1} \exp(-\beta x)$$

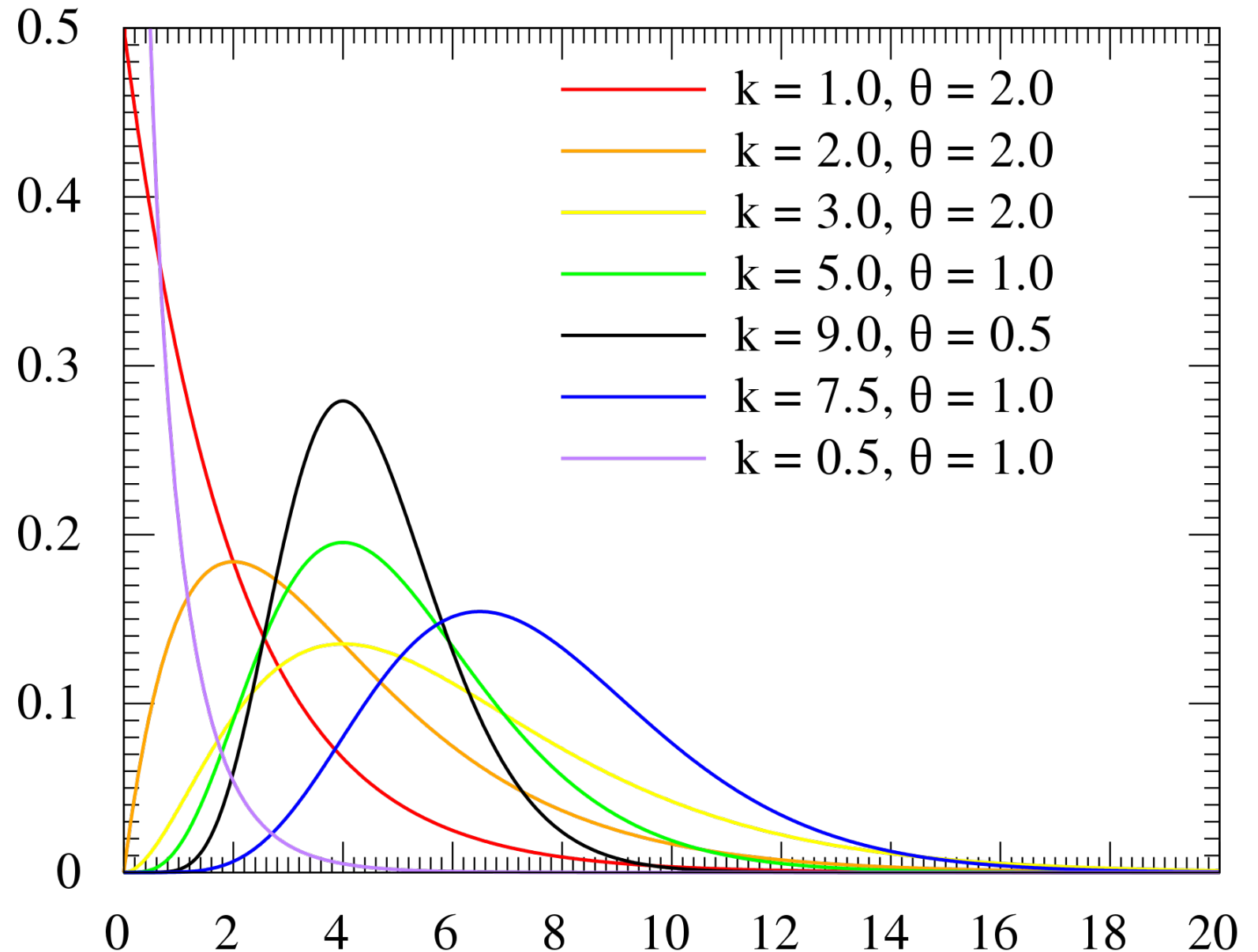
where α (shape) > 0 , β (rate) > 0 , $x > 0$, $\Gamma(\cdot)$ = gamma function defined for all complex numbers with real part positive

- $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx, \quad \Re(z) > 0$
- For positive integer n , $\text{gamma}(n) = \text{factorial}(n-1)$

Transformation of Random Variables

- Gamma PDF

- $k = \text{shape} = \alpha$
- $\text{theta} = \text{scale} = 1/\text{rate} = 1/\beta$
- [Link](#)



Transformation of Random Variables

- Let X have a Gamma PDF,

$$P(x) = \text{Gamma}(x|\alpha, \beta) = (\beta^\alpha / \Gamma(\alpha)) x^{\alpha-1} \exp(-\beta x)$$

where α (shape) > 0 , β (rate) > 0 , $x > 0$, $\Gamma(\cdot)$ = gamma function defined for all complex numbers with real part positive

- $$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx, \quad \Re(z) > 0$$

- For positive integer n , $\text{gamma}(n) = \text{factorial}(n-1)$

- Consider the transformation $Y := 1/X$

- What is the PDF of Y ?

$$y := 1/x \implies x = 1/y \implies g^{-1}(y) = 1/y$$

$$\left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{y^2} \text{ for } y > 0$$

Transformation of Random Variables

- Let X have a Gamma PDF,

$$P(x) = \text{Gamma}(x|\alpha, \beta) = (\beta^\alpha / \Gamma(\alpha)) x^{\alpha-1} \exp(-\beta x)$$

where $\alpha > 0$, $\beta > 0$, $x > 0$, and $\Gamma(\cdot)$ is the well-known gamma function defined for all complex numbers with real part positive.

- Consider the transformation $Y := 1/X$

- What is the PDF of Y ?

$$y := 1/x \implies x = 1/y \implies g^{-1}(y) = 1/y$$

$$\left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{y^2} \text{ for } y > 0$$

$$q_1(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = (\beta^\alpha / \Gamma(\alpha)) y^{1-\alpha} \exp(-\beta/y) \frac{1}{y^2} = (\beta^\alpha / \Gamma(\alpha)) y^{-\alpha-1} \exp(-\beta/y)$$

- This is called the inverse-Gamma PDF

Transformation of Random Variables

- Inverse-Gamma PDF

