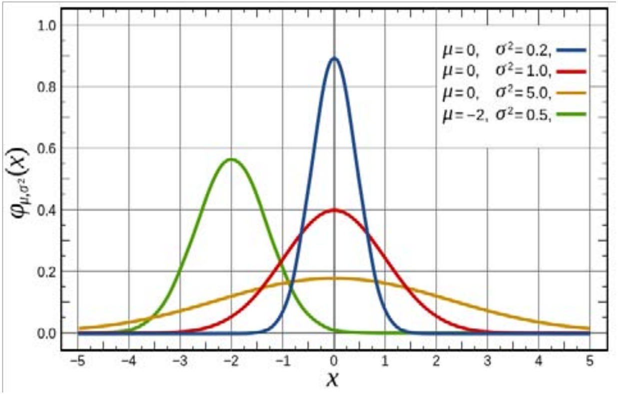
Variance; Continuous Random Variables
18.05 Spring 2014



Variance and standard deviation

X a discrete random variable with mean $E(X) = \mu$.

- Meaning: spread of probability mass about the mean.
- Definition as expectation (weighted sum):

$$\text{Var}(X) = E((X - \mu)^2).$$

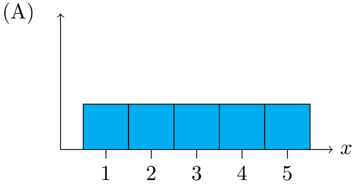
- Computation as sum:

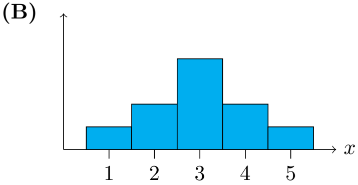
$$\text{Var}(X) = \sum_{i=1}^{n} p(x_i)(x_i - \mu)^2.$$

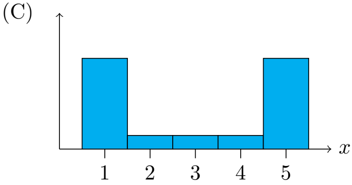
- Standard deviation $\sigma = \sqrt{\text{Var}(X)}$.
  Units for standard deviation = units of X.

Concept question

The graphs below give the pmf for 3 random variables. Order them by size of standard deviation from biggest to smallest. (Assume x has the same units in all 3.)







1. ABC 2. ACB 3. BAC 4. BCA 5. CAB 6. CBA

Computation from tables

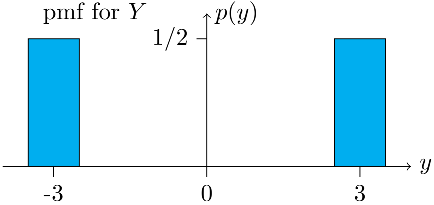
Example. Compute the variance and standard deviation of X.

values x	1	2	3	4	5
pmf $p(x)$	$1/10$	$2/10$	$4/10$	$2/10$	$1/10$

Concept question

Which pmf has the bigger standard deviation? (Assume w and y have the same units.)

1. Y 2. W



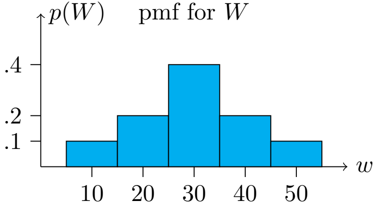


Table question: make probability tables for Y and W and compute their standard deviations.

Concept question

True or false: If $\text{Var}(X) = 0$ then X is constant.

1. True
2. False

Algebra with variances

If a and b are constants then

$$\operatorname{Var}(aX + b) = a^2 \operatorname{Var}(X), \qquad \sigma_{aX+b} = |a|\,\sigma_X.$$

If X and Y are independent random variables then

$$\operatorname{Var}(X + Y) = \operatorname{Var}(X) + \operatorname{Var}(Y).$$

Board questions

1. Prove: if $X \sim \text{Bernoulli}(p)$ then $\text{Var}(X) = p(1-p)$.

2. Prove: if $X \sim \text{bin}(n,p)$ then $\text{Var}(X) = n\,p(1-p)$.

3. Suppose $X_1, X_2, \ldots, X_n$ are independent and all have the same standard deviation $\sigma = 2$. Let $\overline{X}$ be the average of $X_1, \ldots, X_n$.

What is the standard deviation of $\overline{X}$?

Continuous random variables

- Continuous range of values:

$$[0,1],\ [a,b],\ [0,\infty),\ (-\infty,\infty).$$

- Probability density function (pdf)

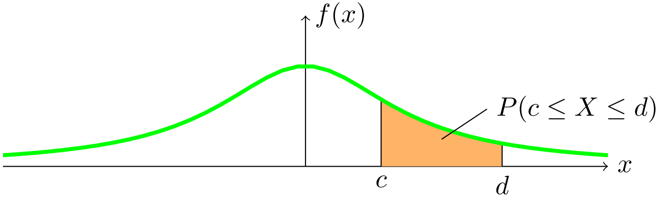
$$f(x) \geq 0; \quad P(c \leq x \leq d) = \int_c^d f(x)\,dx.$$

Units for the pdf are $\dfrac{\text{prob.}}{\text{unit of x}}$

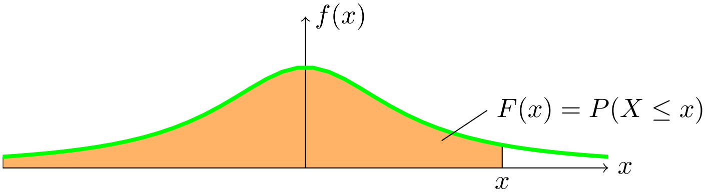
- Cumulative distribution function (cdf)

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt.$$

Visualization



pdf and probability



pdf and cdf

Properties of the cdf

(Same as for discrete distributions)

- (Definition) $F(x) = P(X \leq x)$.
- $0 \leq F(x) \leq 1$.
- non-decreasing.
- 0 to the left: $\lim_{x \to -\infty} F(x) = 0$.
- 1 to the right: $\lim_{x \to \infty} F(x) = 1$.
- $P(c < X \leq d) = F(d) - F(c)$.
- $F'(x) = f(x)$.

Board questions

1. Suppose X has range $[0, 2]$ and pdf $f(x) = cx^2$.

(a) What is the value of c.

(b) Compute the cdf $F(x)$.

(c) Compute $P(1 \leq X \leq 2)$.

2. Suppose Y has range $[0, b]$ and cdf $F(y) = y^2/9$.

(a) What is b?

(b) Find the pdf of Y.

Concept questions

Suppose X is a continuous random variable.

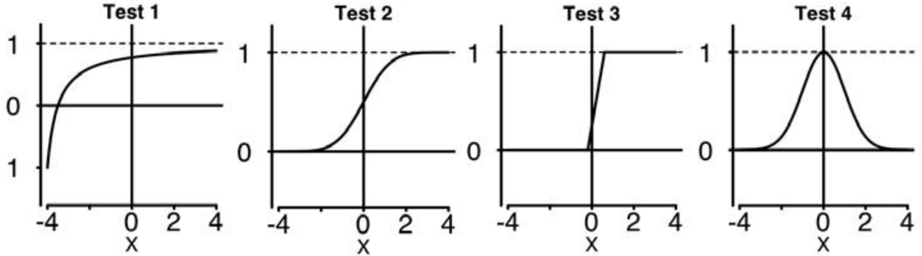
(a) What is $P(a \leq X \leq a)$?

(b) What is $P(X = 0)$?

(c) Does $P(X = a) = 0$ mean X never equals a?

Concept question

Which of the following are graphs of valid cumulative distribution functions?



Add the numbers of the valid cdf's and click that number.

Exponential Random Variables

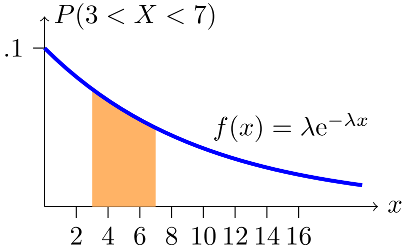
Parameter: λ (called the rate parameter).

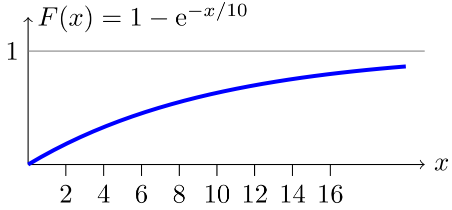
Range: $[0, \infty)$.

Notation: exponential(λ) or exp(λ).

Density: $f(x) = \lambda e^{-\lambda x}$ for $0 \le x$.

Models: Waiting time





Continuous analogue of geometric distribution –memoryless!

Board question

I've noticed that taxis drive past 77 Mass. Ave. on the average of once every 10 minutes.

Suppose time spent waiting for a taxi is modeled by an exponential random variable

$$X \sim \text{Exponential}(1/10); \qquad f(x) = \frac{1}{10}\mathrm{e}^{-x/10}$$

(a) Sketch the pdf of this distribution

(b) Shade the region which represents the probability of waiting between 3 and 7 minutes

(c) Compute the probability of waiting between between 3 and 7 minutes for a taxi

(d) Compute and sketch the cdf.