Continuous r.v.'s: cdf's, Expected Values

Engineering Statistics
Section 4.2

Josh Engwer

TTU

29 February 2016

PART I:

CUMULATIVE DENSITY FUNCTION (CDF) OF A CONTINUOUS RANDOM VARIABLE

Cumulative Density Fcn (cdf) of a Continuous r.v.

Definition

(cdf of a Continuous Random Variable)

Let X be a **continuous** random variable with pdf $f_X(x)$.
Then, its **cdf**, denoted as $F_X(x)$, is defined as follows:

$$F_X(x) := \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(\xi)\, d\xi = \int_{-\infty}^{x} f_X(t)\, dt$$

Corollary

(cdf Axioms)

*Let X be a **continuous** random variable. Then, its **cdf** $F_X(x)$, satisfies*

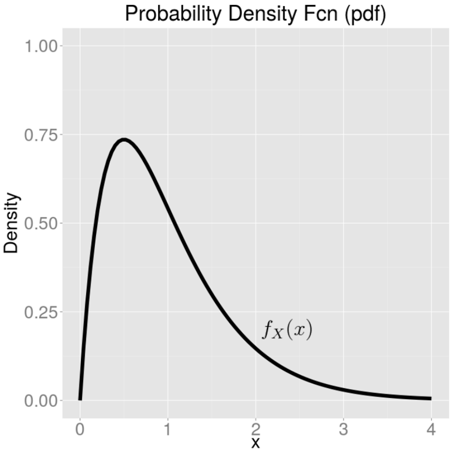
Eventually Zero (One) to the Left (Right): $\lim_{x \to -\infty} F_X(x) = 0, \quad \lim_{x \to +\infty} F_X(x) = 1$

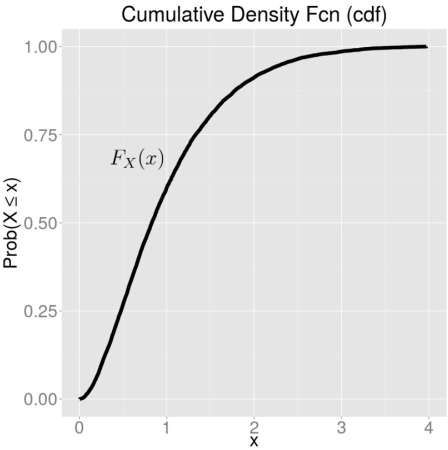
Non-decreasing: $x_1 \leq x_2 \implies F_X(x_1) \leq F_X(x_2)$

Continuous: $F_X \in C(\mathbb{R})$

<u>NOTATION:</u> ξ is the lowercase Greek letter "xi", pronounced (kuh)-SEE.

Plots of pdf & cdf of a continuous r.v.





Computing Probabilities using a Continuous cdf

Theorem

*Let X be a **continuous** r.v. with cdf $F_X(x)$. Let scalars $a, b \in \mathbb{R}$ s.t. $a < b$. Then:*

$$\begin{aligned}
\mathbb{P}(X \le a) &= \mathbb{P}(X < a) &&= F_X(a) \\
\mathbb{P}(a \le X \le b) &= \mathbb{P}(a < X < b) &&= F_X(b) - F_X(a) \\
\mathbb{P}(a < X \le b) &= \mathbb{P}(a \le X < b) &&= F_X(b) - F_X(a) \\
\mathbb{P}(X \ge b) &= \mathbb{P}(X > b) &&= 1 - F_X(b)
\end{aligned}$$

Proposition

(Obtaining the pdf from a cdf of a continuous r.v.)

*Let X be a **continuous** r.v. with pdf $f_X(x)$ and cdf $F_X(x)$. Then:* $F'_X(x) = f_X(x)$

PART II:

EXPECTED VALUE, VARIANCE, STANDARD DEVIATION OF A CONTINUOUS RANDOM VARIABLE

Expected Value (Mean) of a Continuous r.v.

Definition

(Expected Value of a Continuous r.v.)

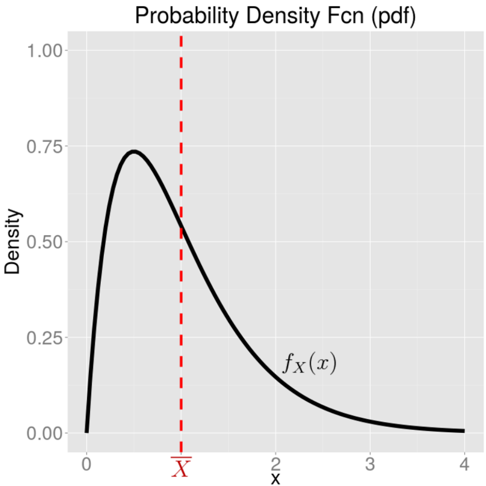
Let X be a **continuous** random variable with pdf $f_X(x)$.
Then the **expected value** (AKA **mean**) of X is:

$$\mathbb{E}[X] := \int_{\text{Supp}(X)} x \cdot f_X(x)\ dx$$

It's possible (but rare) that the expected value is **infinite**: $\mathbb{E}[X] = \infty$

NOTATION: The expected value of X is alternatively denoted by $\overline{X}$ or μ_X.

Mean of a Continuous Distribution (Example)



The mean of a continuous distribution is the location of the fulcrum on the horizontal axis of pdf which balances the data.

In this particular case, $\overline{X} = 1$. (dashed red line)

Expected Value (Mean) of Continuous Function of r.v.

Definition

Let X be a **continuous** random variable with pdf $f_X(x)$.
Let $h(x)$ be a single-variable function.
Then the **expected value** (AKA **mean**) of $h(X)$ is:

$$\mathbb{E}[h(X)] := \int_{\text{Supp}(X)} h(x) \cdot f_X(x)\, dx$$

It's possible (but rare) that the expected value is **infinite**: $\mathbb{E}[h(X)] = \pm\infty$

NOTATION: The expected value of $h(X)$ is alternatively denoted by $\mu_{h(X)}$.

Variance & Standard Deviation of a Continuous r.v.

Definition

(Variance & Standard Deviation of a Continuous Random Variable)

Let X be a **continuous** random variable with pdf $f_X(x)$ and mean μ_X.
Then the **variance** of X is:

$$\mathbb{V}[X] := \mathbb{E}[(X - \mu_X)^2] = \int_{\text{Supp}(X)} (x - \mu_X)^2 \cdot f_X(x)\, dx$$

Moreover, the **standard deviation** of X is: $\sigma_X := \sqrt{\mathbb{V}[X]}$

It's possible (but rare) that the variance is **infinite**: $\mathbb{V}[X] = \infty$

<u>NOTATION:</u> The variance of X is alternatively denoted by σ_X^2 or $\text{Var}(X)$.

An Easier Way to Compute Variance

Computing variances using the definition can be quite tedious! Fortunately, there's an equivalent formula that's easier to use:

Corollary

(Easier Formula for Variance)

*Let X be a **continuous** random variable with pdf $f_X(x)$. Then:*

$$\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

<u>PROOF:</u> Let $\mu_X \equiv \mathbb{E}[X]$. Then:

$$\begin{aligned}
\mathbb{V}[X] &:= \int_{\text{Supp}(X)} (x-\mu_X)^2 \cdot f_X(x)\,dx = \int_{\text{Supp}(X)} (x^2 - 2x\mu_X + \mu_X^2) \cdot f_X(x)\,dx \\
&= \int_{\text{Supp}(X)} x^2 \cdot f_X(x)\,dx - 2\mu_X \int_{\text{Supp}(X)} x \cdot f_X(x)\,dx + \mu_X^2 \int_{\text{Supp}(X)} f_X(x)\,dx \\
&:= \mathbb{E}[X^2] - 2\mu_X \mathbb{E}[X] + \mu_X^2 \cdot 1 = \mathbb{E}[X^2] - 2\mu_X(\mu_X) + \mu_X^2 \\
&= \mathbb{E}[X^2] - 2\mu_X^2 + \mu_X^2 = \mathbb{E}[X^2] - \mu_X^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2
\end{aligned}$$

QED

PART III:

QUANTILES/PERCENTILES, QUARTILES, MEDIAN
OF A CONTINUOUS DISTRIBUTION

Quantiles/Percentiles & Median of a Continuous Dist.

Definition

(Quantile/Percentile & Median of a Continuous Distribution)

Let X be a **continuous** r.v. with pdf $f_X(x)$ & cdf $F_X(x)$, and let $0 < p < 1$. Then:

(1) $x_p \equiv$ The p**-quantile** of the distribution of X and is defined by

$$p = F_X(x_p) \iff p = \int_{-\infty}^{x_p} f_X(\xi)\, d\xi = \int_{-\infty}^{x_p} f_X(t)\, dt$$

i.e. Solve equation $p = F_X(x_p)$ for x_p.

(2) $x_p \equiv$ The $(100p)$**-th percentile** of the distribution of X.

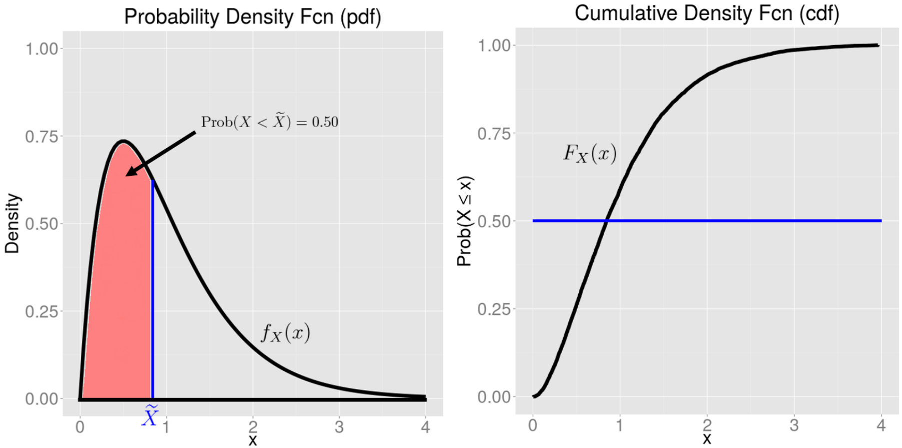
e.g. $x_{0.37} \equiv$ The 0.37-quantile (37^{th} percentile) of the distribution of X.

(3) $x_{0.25} \equiv$ The 1^{st} **quartile** (25^{th} percentile) of the distribution of X.

(4) $\widetilde{X} = x_{0.50} \equiv$ The **median** (50^{th} percentile) of the distribution of X.

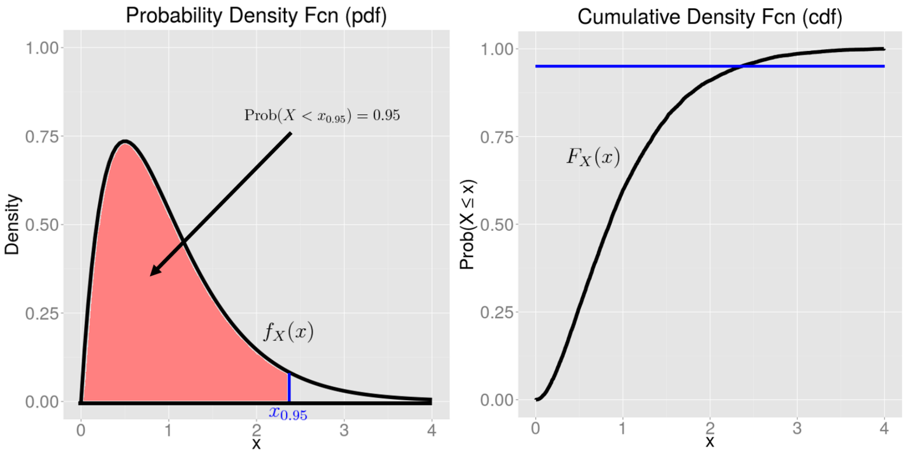
(5) $x_{0.75} \equiv$ The 3^{rd} **quartile** (75^{th} percentile) of the distribution of X.

Median of Continuous Distribution (Example)



When viewing the pdf of a continuous distribution (left plot), 50% of the data is to the left of the median. (blue line)

95^{th} Percentile of Continuous Distribution (Example)



When viewing the pdf of a continuous distribution (left plot), 95% of the data is to the left of the 95^{th} percentile. (blue line)

Textbook Logistics for Section 4.2

- Difference(s) in Notation:

CONCEPT	TEXTBOOK NOTATION	SLIDES/OUTLINE NOTATION
Probability of Event	$P(A)$	$\mathbb{P}(A)$
Support of a r.v.	"All possible values of X"	Supp(X)
pdf of a r.v.	$f(x)$	$f_X(x)$
cdf of a r.v.	$F(x)$	$F_X(x)$
Expected Value of r.v.	$E(X)$	$\mathbb{E}[X]$
Variance of r.v.	$V(X)$	$\mathbb{V}[X]$
Median of r.v.	$\widetilde{\mu}$	$\widetilde{\mu}_X$
$(100p)^{th}$ Percentile of r.v.	$\eta(p)$	x_p

NOTE: η is the lower-case Greek letter "eta"

Fin

Fin.