MATH E-156 Mathematical Statistics

Harvard Extension School

Dmitry Kurochkin

Fall 2020
Lecture 5

Contents

Contents

Uniform

$X \sim \text{Uniform}[a, b]$, where $a < b$, if X is a continuous random variable with the following probability density function (pdf):

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{for all } x \in [a, b], \\ 0, & \text{otherwise.} \end{cases}$$

<u>Claim</u>:

If $X \sim \text{Uniform}[a, b]$ with some $b > 0$ then

1. the cumulative distribution function (cdf) is

$$F_X(x) = \begin{cases} 0, & \text{if } x < a, \\ \frac{x-a}{b-a}, & \text{for all } x \in [a, b], \\ 1, & \text{if } x > b; \end{cases}$$

2. the expectation is

$$E[X] = \frac{a+b}{2}.$$

Contents

Exponential

$X \sim \text{Exponential}(\lambda)$, where $\lambda > 0$, if X is a continuous random variable with the following probability density function (pdf):

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{for all } x \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

<u>Claim</u>:

If $X \sim \text{Exponential}(\lambda)$ with some $\lambda > 0$ then

1. the cumulative distribution function (cdf) is

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{for all } x \geq 0, \\ 0, & \text{otherwise;} \end{cases}$$

2. the expectation is

$$E[X] = \frac{1}{\lambda}.$$

Contents

Normal

$X \sim \text{Normal}(\mu, \sigma^2)$, where $\mu \in \mathbb{R}$ and $\sigma > 0$, if X is a continuous random variable with the following probability density function (pdf):

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{for all } x \in \mathbb{R}.$$

Claim:

If $X \sim \text{Normal}(\mu, \sigma^2)$ with some $\mu \in \mathbb{R}$ and $\sigma > 0$ then

1. the cumulative distribution function (cdf) is

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right), \quad \text{where } \Phi(z) \doteq \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt;$$

2. the expectation is

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

Contents

Gamma

$X \sim$ Gamma(α, λ), where $\alpha, \lambda > 0$, if X is a continuous random variable with the following probability density function (pdf):

$$f_X(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & \text{for all } x \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Here,

$$\Gamma(x) \doteq \int_0^{+\infty} u^{x-1} e^{-u} du \;\; \text{for} \;\; x > 0.$$

<u>Claim</u>:
If $X \sim$ Gamma(α, λ) with some $\alpha, \lambda > 0$ then

$$E[X] = \frac{\alpha}{\lambda}.$$

Contents

Sum of Discrete Random Variables

Claim:

Let X and Y be two discrete random variables that take integer values. Define a new random variable as

$$S = X + Y$$

then the probability mass function (pmf) of S is

$$p_S(s) = \sum_{x=-\infty}^{+\infty} p_{X,Y}(x, s-x) \quad \text{for all integer } s.$$

Contents

Sum of Continuous Random Variables

Claim:

Let X and Y be two continuous random variables that take integer values. Define a new random variable as

$$S = X + Y$$

then the probability density function (pdf) of S is

$$f_S(s) = \int_{-\infty}^{+\infty} f_{X,Y}(x, s - x)dx \quad \text{for all integer } s.$$

Contents

Order Statistics

Def.

Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables. Then we define

- the largest value

$$X_{(n)} \doteq \max\{X_1, X_2, \ldots, X_n\};$$

Order Statistics

Def.

Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables. Then we define

- the largest value

$$X_{(n)} \doteq \max\{X_1, X_2, \ldots, X_n\};$$

- the smallest value

$$X_{(1)} \doteq \min\{X_1, X_2, \ldots, X_n\};$$

Order Statistics

Def.

Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables. Then we define

- the largest value

$$X_{(n)} \doteq \max\{X_1, X_2, \ldots, X_n\};$$

- the smallest value

$$X_{(1)} \doteq \min\{X_1, X_2, \ldots, X_n\};$$

- for each $k \in \{1, 2, \ldots, n\}$, we denote the kth-smallest value of $X_1, X_2, \ldots, X_n$ by let $X_{(k)}$.

  Note:

  For any sample $X_1, X_2, \ldots, X_n$, we have:

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}.$$

Order Statistics

Claim:
Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables with the common cdf $F(x)$ and pdf $f(x)$. Then

1. the cdf of $X_{(n)}$ is
$$F_{X_{(n)}}(x) = [F(x)]^n;$$
2. the pdf of $X_{(n)}$ is
$$f_{X_{(n)}}(x) = nf(x)[F(x)]^{n-1}.$$

Order Statistics

Claim:

Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables with the common cdf $F(x)$ and pdf $f(x)$. Then

1. the cdf of $X_{(1)}$ is
$$F_{X_{(1)}}(x) = 1 - [1 - F(x)]^n;$$
2. the pdf of $X_{(1)}$ is
$$f_{X_{(1)}}(x) = nf(x)[1 - F(x)]^{n-1}.$$

Order Statistics

Claim:
Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables with the common cdf $F(x)$ and pdf $f(x)$.
Then for each $k \in \{1, 2, \ldots, n\}$, the pdf of $X_{(k)}$ is

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!} f(x)[F(x)]^{k-1}[1 - F(x)]^{n-k}.$$

Contents

Expectations of Functions of Random Variables

THEOREM A

Suppose that $Y = g(X)$.

a. If X is discrete with frequency function $p(x)$, then

$$E(Y) = \sum_x g(x)p(x)$$

provided that $\sum |g(x)|p(x) < \infty$.

b. If X is continuous with density function $f(x)$, then

$$E(Y) = \int_{-\infty}^{\infty} g(x)f(x)\,dx$$

provided that $\int |g(x)|f(x)\,dx < \infty$.

Source: *Mathematical Statistics and Data Analysis* by John Rice

Expectations of Functions of Random Variables

THEOREM B

Suppose that $X_1, \ldots, X_n$ are jointly distributed random variables and $Y = g(X_1, \ldots, X_n)$.

a. If the X_i are discrete with frequency function $p(x_1, \ldots, x_n)$, then

$$E(Y) = \sum_{x_1, \ldots, x_n} g(x_1, \ldots, x_n) p(x_1, \ldots, x_n)$$

provided that $\sum_{x_1, \ldots, x_n} |g(x_1, \ldots, x_n)| p(x_1, \ldots, x_n) < \infty$.

b. If the X_i are continuous with joint density function $f(x_1, \ldots, x_n)$, then

$$E(Y) = \int \int \cdots \int g(x_1, \ldots, x_n) f(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

provided that the integral with $|g|$ in place of g converges.

Source: *Mathematical Statistics and Data Analysis* by John Rice

Contents

Expectations of Functions of Random Variables

Claim:
Let $X_1, X_2, \ldots, X_n$ be jointly distributed random variables with expectations $\mathrm{E}[X_1], \mathrm{E}[X_2], \ldots, \mathrm{E}[X_n]$, respectively.
Then

$$\mathrm{E}\left[a + \sum_{k=1}^{n} b_k X_k\right] = a + \sum_{k=1}^{n} b_k \mathrm{E}[X_k],$$

for any constants $a, b_1, b_2, \ldots, b_n \in \mathbb{R}$.

Contents

Variance and Standard Deviation

Def.

Let X be a random variable. Then, provided the expectations exist,

- Variance of X is defined as

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

- Note that $\mathrm{Var}(X) \geq 0$. Standard deviation of X is defined as

$$\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}.$$

Contents

Properties of Variance

Let X be a random variable for which $\mathrm{Var}(X)$ exists and $a, b \in \mathbb{R}$ be some constants, then

1. $\mathrm{Var}(X) \geq 0$
2. $\mathrm{Var}(a) = 0$
3. $\mathrm{Var}(X) = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$
4. $\mathrm{Var}(a + bX) = b^2 \mathrm{Var}(X)$
5. $\mathrm{SD}(a + bX) = |b| \mathrm{SD}(X)$

Variance of a Discrete Random Variable

Examples:

1. Let $X \sim \text{Bernoilli}(p)$ with $p \in [0,1]$ then
$$\mathrm{E}[X] = p \quad \text{and} \quad \mathrm{Var}(X) = p(1-p).$$

2. Let $X \sim \text{Binomial}(n,p)$ with $n \in \{1,2,3,\ldots\}$ and $p \in [0,1]$ then
$$\mathrm{E}[X] = np \quad \text{and} \quad \mathrm{Var}(X) = np(1-p).$$

3. Let $X \sim \text{Geometric}(p)$ with $p \in (0,1]$ then
$$E[X] = \frac{1}{p} \quad \text{and} \quad \mathrm{Var}(X) = \frac{1-p}{p^2}.$$

4. Let X take $1,2,3,\ldots$ values with $P(X=k) = \frac{6}{\pi^2}\frac{1}{k^2}$ for all $k \in \{1,2,3,\ldots\}$ then
$$E[X] \ \text{ is not defined; and thus } \ \mathrm{Var}(X) \text{ is not defined.}$$

Variance of a Continuous Random Variable

Examples:

1. Let $X \sim \text{Uniform}[a,b]$ with $a < b$, that is,

$$f_X(x) = \begin{cases} 0, & \text{if } x < a, \\ \frac{1}{b-a}, & \text{if } a \leq x \leq b, \\ 0, & \text{if } x > b, \end{cases}$$

then

$$E[X] = \frac{a+b}{2} \quad \text{and} \quad \mathrm{Var}(X) = \frac{(b-a)^2}{12}.$$

2. Let $X \sim N(0, 1^2)$, i.e. be *standard normal*, that is,

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \text{for all } x \in \mathbb{R},$$

then

$$E[X] = 0 \quad \text{and} \quad \mathrm{Var}(X) = 1.$$

Contents

Chebyshev's Inequality

<u>Thm.</u> (Chebyshev's Inequality)
Let X be a random variable with mean $\mathrm{E}[X] = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$. Then, for any $t > 0$,

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}.$$

Chebyshev's Inequality

Thm. (Chebyshev's Inequality)
Let X be a random variable with mean $\mathrm{E}[X] = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$. Then, for any $t > 0$,

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}.$$

Note:
For $k > 0$, let $t = k\sigma$ then

$$P(|X - \mu| > k\sigma) \leq \frac{1}{k}.$$

Corollary:
If $\mathrm{Var}(X) = 0$, then $P(X = \mu) = 1$.

Contents

Bias and Mean Squared Error

Def.

Let X be a measurement (random variable) of some true value x_0. Then

- *Bias* is defined as
$$\text{Bias} = \mathrm{E}[X - x_0].$$
- *Mean Squared Error* (MSE) is defined as
$$\text{MSE} = \mathrm{E}\left[(X - x_0)^2\right].$$

Bias and Mean Squared Error

Def.

Let X be a measurement (random variable) of some true value x_0. Then

- *Bias* is defined as
$$\text{Bias} = \mathrm{E}[X - x_0].$$
- *Mean Squared Error* (MSE) is defined as
$$\text{MSE} = \mathrm{E}\left[(X - x_0)^2\right].$$

Thm.

Let X be a measurement of some value x_0. Then
$$\text{MSE} = \text{Bias}^2 + \mathrm{Var}(X).$$