EKT-816 Lecture 1

Probability Review (1)

Jesse Naidoo

University of Pretoria

PDFs, CDFs, and Quantiles

- Discrete distribution:
  - mass functions: $f(x) = P(X = x)$.
  - cumulative distribution functions: $F(x) = P(X \leq x)$.
  - Examples: Bernoulli(p); binomial(n, p); Poisson(λ).
- continuous distributions:
  - density function $f_X(x)$ such that

$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$

  - CDF $F_X(x)$ is increasing and such that $\lim_{x \to -\infty} F_X(x) = 0$, $\lim_{x \to \infty} F_X(x) = 1$.
  - $F_X'(x) = f_X(x)$, or

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

- the τ-th quantile of the distribution F is the value x_τ such that

$$F(x_\tau) = \tau$$

PDFs, CDFs, and Quantiles

- Discrete distribution:
  - mass functions: $f(x) = P(X = x)$.
  - cumulative distribution functions: $F(x) = P(X \leq x)$.
  - Examples: Bernoulli(p); binomial(n, p); Poisson(λ).
- continuous distributions:
  - density function $f_X(x)$ such that

$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$

  - CDF $F_X(x)$ is increasing and such that $\lim_{x \to -\infty} F_X(x) = 0$, $\lim_{x \to \infty} F_X(x) = 1$.
  - $F_X'(x) = f_X(x)$, or

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

- the τ-th quantile of the distribution F is the value x_τ such that

$$F(x_\tau) = \tau$$

PDFs, CDFs, and Quantiles

- Discrete distribution:
  - mass functions: $f(x) = P(X = x)$.
  - cumulative distribution functions: $F(x) = P(X \le x)$.
  - Examples: Bernoulli(p); binomial(n, p); Poisson(λ).
- continuous distributions:
  - density function $f_X(x)$ such that

$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$

  - CDF $F_X(x)$ is increasing and such that $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to\infty} F_X(x) = 1$.
  - $F'_X(x) = f_X(x)$, or

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

- the τ-th quantile of the distribution F is the value x_τ such that

$$F(x_\tau) = \tau$$

PDFs, CDFs, and Quantiles

- Discrete distribution:
  - mass functions: $f(x) = P(X = x)$.
  - cumulative distribution functions: $F(x) = P(X \leq x)$.
  - Examples: Bernoulli(p); binomial(n, p); Poisson(λ).
- continuous distributions:
  - density function $f_X(x)$ such that

$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$

  - CDF $F_X(x)$ is increasing and such that $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to\infty} F_X(x) = 1$.
  - $F_X'(x) = f_X(x)$, or

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

- the τ-th quantile of the distribution F is the value x_τ such that

$$F(x_\tau) = \tau$$

PDFs, CDFs, and Quantiles

- Discrete distribution:
  - mass functions: $f(x) = P(X = x)$.
  - cumulative distribution functions: $F(x) = P(X \leq x)$.
  - Examples: Bernoulli(p); binomial(n, p); Poisson(λ).
- continuous distributions:
  - density function $f_X(x)$ such that
$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$
  - CDF $F_X(x)$ is increasing and such that $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to\infty} F_X(x) = 1$.
  - $F_X'(x) = f_X(x)$, or
$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$
- the τ-th quantile of the distribution F is the value x_τ such that
$$F(x_\tau) = \tau$$

PDFs, CDFs, and Quantiles

- Discrete distribution:
    - mass functions: $f(x) = P(X = x)$.
    - cumulative distribution functions: $F(x) = P(X \leq x)$.
    - Examples: Bernoulli(p); binomial(n, p); Poisson(λ).
- continuous distributions:
    - density function $f_X(x)$ such that

$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$

    - CDF $F_X(x)$ is increasing and such that $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to\infty} F_X(x) = 1$.
    - $F_X'(x) = f_X(x)$, or

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

- the τ-th quantile of the distribution F is the value x_τ such that

$$F(x_\tau) = \tau$$

PDFs, CDFs, and Quantiles

- Discrete distribution:
  - mass functions: $f(x) = P(X = x)$.
  - cumulative distribution functions: $F(x) = P(X \leq x)$.
  - Examples: Bernoulli(p); binomial(n, p); Poisson(λ).
- continuous distributions:
  - density function $f_X(x)$ such that

$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$

  - CDF $F_X(x)$ is increasing and such that $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to\infty} F_X(x) = 1$.
  - $F'_X(x) = f_X(x)$, or

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

- the τ-th quantile of the distribution F is the value x_τ such that

$$F(x_\tau) = \tau$$

PDFs, CDFs, and Quantiles

- Discrete distribution:
  - mass functions: $f(x) = P(X = x)$.
  - cumulative distribution functions: $F(x) = P(X \leq x)$.
  - Examples: Bernoulli(p); binomial(n, p); Poisson(λ).
- continuous distributions:
  - density function $f_X(x)$ such that

$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$

  - CDF $F_X(x)$ is increasing and such that $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to\infty} F_X(x) = 1$.
  - $F_X'(x) = f_X(x)$, or

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

- the τ-th quantile of the distribution F is the value x_τ such that

$$F(x_\tau) = \tau$$

PDFs, CDFs, and Quantiles

- Discrete distribution:
  - mass functions: $f(x) = P(X = x)$.
  - cumulative distribution functions: $F(x) = P(X \leq x)$.
  - Examples: Bernoulli(p); binomial(n, p); Poisson(λ).
- continuous distributions:
  - density function $f_X(x)$ such that

$$\int_{-\infty}^{\infty} f_X(x)dx = 1.$$

  - CDF $F_X(x)$ is increasing and such that $\lim_{x\to-\infty} F_X(x) = 0$, $\lim_{x\to\infty} F_X(x) = 1$.
  - $F'_X(x) = f_X(x)$, or

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

- the τ-th quantile of the distribution F is the value x_τ such that

$$F(x_\tau) = \tau$$

Moments

- the *mean* of a distribution is

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

- the *variance* of the distribution is

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

  where μ is the mean of the distribution
- note, these moments may not exist!
  - but, if $V[X] < \infty$, the mean will exist (why?)
  - also notice that $V[X] = E[X^2] - E[X]^2$
  - third (centered) moment is called *skewness*
  - fourth (centered) moment is called *kurtosis*

Moments

- the *mean* of a distribution is

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

- the *variance* of the distribution is

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

  where μ is the mean of the distribution
- note, these moments may not exist!
  - but, if $V[X] < \infty$, the mean will exist (why?)
  - also notice that $V[X] = E[X^2] - E[X]^2$
  - third (centered) moment is called *skewness*
  - fourth (centered) moment is called *kurtosis*

Moments

- the *mean* of a distribution is

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

- the *variance* of the distribution is

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

  where μ is the mean of the distribution
- note, these moments may not exist!
  - but, if $V[X] < \infty$, the mean will exist (why?)
  - also notice that $V[X] = E[X^2] - E[X]^2$
  - third (centered) moment is called *skewness*
  - fourth (centered) moment is called *kurtosis*

Moments

- the *mean* of a distribution is

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

- the *variance* of the distribution is

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

  where μ is the mean of the distribution
- note, these moments may not exist!
  - but, if $V[X] < \infty$, the mean will exist (why?)
  - also notice that $V[X] = E[X^2] - E[X]^2$
  - third (centered) moment is called *skewness*
  - fourth (centered) moment is called *kurtosis*

Moments

- the *mean* of a distribution is

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

- the *variance* of the distribution is

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

  where μ is the mean of the distribution
- note, these moments may not exist!
  - but, if $V[X] < \infty$, the mean will exist (why?)
  - also notice that $V[X] = E[X^2] - E[X]^2$
  - third (centered) moment is called *skewness*
  - fourth (centered) moment is called *kurtosis*

Moments

- the *mean* of a distribution is

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

- the *variance* of the distribution is

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

  where μ is the mean of the distribution
- note, these moments may not exist!
  - but, if $V[X] < \infty$, the mean will exist (why?)
  - also notice that $V[X] = E[X^2] - E[X]^2$
  - third (centered) moment is called *skewness*
  - fourth (centered) moment is called *kurtosis*

Moments

- the *mean* of a distribution is

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)dx$$

- the *variance* of the distribution is

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

  where μ is the mean of the distribution
- note, these moments may not exist!
  - but, if $V[X] < \infty$, the mean will exist (why?)
  - also notice that $V[X] = E[X^2] - E[X]^2$
  - third (centered) moment is called *skewness*
  - fourth (centered) moment is called *kurtosis*

Example: Pareto Distributions

- let $\alpha > 0$ be some constant
- density is

$$f_X(x) = \begin{cases} \alpha x^{-(\alpha+1)} & \text{if } x > 1 \\ 0 & \text{else} \end{cases}$$

  - what is the CDF, $F_X(x)$?
  - what is the mean, $E[X]$? do we have to impose any conditions to ensure the mean exists?
  - what is the variance, $V[X]$?

Example: Pareto Distributions

- let $\alpha > 0$ be some constant
- density is

$$f_X(x) = \begin{cases} \alpha x^{-(\alpha+1)} & \text{if } x > 1 \\ 0 & \text{else} \end{cases}$$

  - what is the CDF, $F_X(x)$?
  - what is the mean, $E[X]$? do we have to impose any conditions to ensure the mean exists?
  - what is the variance, $V[X]$?

Example: Pareto Distributions

- let $\alpha > 0$ be some constant
- density is

$$f_X(x) = \begin{cases} \alpha x^{-(\alpha+1)} & \text{if } x > 1 \\ 0 & \text{else} \end{cases}$$

  - what is the CDF, $F_X(x)$?
  - what is the mean, $E[X]$? do we have to impose any conditions to ensure the mean exists?
  - what is the variance, $V[X]$?

Example: Pareto Distributions

- let $\alpha > 0$ be some constant
- density is

$$f_X(x) = \begin{cases} \alpha x^{-(\alpha+1)} & \text{if } x > 1 \\ 0 & \text{else} \end{cases}$$

  - what is the CDF, $F_X(x)$?
  - what is the mean, $E[X]$? do we have to impose any conditions to ensure the mean exists?
  - what is the variance, $V[X]$?

Example: Pareto Distributions

- let $\alpha > 0$ be some constant
- density is

$$f_X(x) = \begin{cases} \alpha x^{-(\alpha+1)} & \text{if } x > 1 \\ 0 & \text{else} \end{cases}$$

  - what is the CDF, $F_X(x)$?
  - what is the mean, $E[X]$? do we have to impose any conditions to ensure the mean exists?
  - what is the variance, $V[X]$?

Inverse CDF Trick

- suppose we want to generate random numbers from some distribution with CDF F
  - we can compute F and F^{-1}
  - we can generate uniformly distributed random numbers, $U \sim U(0, 1)$
- then, you can generate $X \sim F$ as follows:

$$X = F^{-1}(U)$$

- proof: let's find the CDF of such Xs.
  - let x be an arbitrary number; we're going to show that $P(X \leq x) = F(x)$
  - $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$

Inverse CDF Trick

- suppose we want to generate random numbers from some distribution with CDF F
  - we can compute F and F^{-1}
  - we can generate uniformly distributed random numbers, $U \sim U(0,1)$
- then, you can generate $X \sim F$ as follows:

$$X = F^{-1}(U)$$

- proof: let's find the CDF of such Xs.
  - let x be an arbitrary number; we're going to show that $P(X \leq x) = F(x)$
  - $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$

Inverse CDF Trick

- suppose we want to generate random numbers from some distribution with CDF F
  - we can compute F and F^{-1}
  - we can generate uniformly distributed random numbers, $U \sim U(0, 1)$
- then, you can generate $X \sim F$ as follows:

$$X = F^{-1}(U)$$

- proof: let's find the CDF of such Xs.
  - let x be an arbitrary number; we're going to show that $P(X \leq x) = F(x)$
  - $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$

Inverse CDF Trick

- suppose we want to generate random numbers from some distribution with CDF F
  - we can compute F and F^{-1}
  - we can generate uniformly distributed random numbers, $U \sim U(0, 1)$
- then, you can generate $X \sim F$ as follows:

$$X = F^{-1}(U)$$

- proof: let's find the CDF of such Xs.
  - let x be an arbitrary number; we're going to show that $P(X \leq x) = F(x)$
  - $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$

Inverse CDF Trick

- suppose we want to generate random numbers from some distribution with CDF F
  - we can compute F and F^{-1}
  - we can generate uniformly distributed random numbers, $U \sim U(0,1)$
- then, you can generate $X \sim F$ as follows:

$$X = F^{-1}(U)$$

- proof: let's find the CDF of such Xs.
  - let x be an arbitrary number; we're going to show that $P(X \leq x) = F(x)$
  - $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$

Inverse CDF Trick

- suppose we want to generate random numbers from some distribution with CDF F
  - we can compute F and F^{-1}
  - we can generate uniformly distributed random numbers, $U \sim U(0,1)$
- then, you can generate $X \sim F$ as follows:

$$X = F^{-1}(U)$$

- proof: let's find the CDF of such Xs.
  - let x be an arbitrary number; we're going to show that $P(X \leq x) = F(x)$
  - $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$

Inverse CDF Trick

- suppose we want to generate random numbers from some distribution with CDF F
  - we can compute F and F^{-1}
  - we can generate uniformly distributed random numbers, $U \sim U(0, 1)$
- then, you can generate $X \sim F$ as follows:

$$X = F^{-1}(U)$$

- proof: let's find the CDF of such Xs.
  - let x be an arbitrary number; we're going to show that $P(X \leq x) = F(x)$
  - $P(X \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x)$

Marginal and Conditional Distributions

- take a joint density $f_{XY}(x, y)$ that integrates to 1 over $\mathbb{R}^2$
  - the marginal density of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy$$

  - analogous for marginal of Y
- the *conditional* density of Y *given that* $X = x$ is

$$f_{Y|X}(y|X = x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

Marginal and Conditional Distributions

- take a joint density $f_{XY}(x, y)$ that integrates to 1 over $\mathbb{R}^2$
  - the marginal density of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$$

  - analogous for marginal of Y
- the *conditional* density of Y *given that* $X = x$ is

$$f_{Y|X}(y|X = x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

Marginal and Conditional Distributions

- take a joint density $f_{XY}(x, y)$ that integrates to 1 over $\mathbb{R}^2$
  - the marginal density of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy$$

  - analogous for marginal of Y
- the *conditional* density of Y *given that* $X = x$ is

$$f_{Y|X}(y|X = x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

Marginal and Conditional Distributions

- take a joint density $f_{XY}(x, y)$ that integrates to 1 over $\mathbb{R}^2$
  - the marginal density of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)dy$$

  - analogous for marginal of Y
- the *conditional* density of Y *given that* $X = x$ is

$$f_{Y|X}(y|X = x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

Basic Rules

- expectations are linear: $E[aX + Y] = aE[X] + E[Y]$
- $V[aX] = a^2 V[X]$
- $V[X + Y] = V[X] + V[Y] + 2\,\mathrm{cov}(X, Y)$

Basic Rules

- expectations are linear: $E[aX + Y] = aE[X] + E[Y]$
- $V[aX] = a^2 V[X]$
- $V[X + Y] = V[X] + V[Y] + 2\,\mathrm{cov}(X, Y)$

Basic Rules

- expectations are linear: $E[aX + Y] = aE[X] + E[Y]$
- $V[aX] = a^2 V[X]$
- $V[X + Y] = V[X] + V[Y] + 2\,\mathrm{cov}(X, Y)$

Example: Zero Correlation, But Not Independent

- consider the following distribution:

$$f_{XY}(x,y) = \begin{cases} 3/4 & \text{if } x \in (-1,1) \text{ and } y \in (0, 1-x^2) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

- show that $\mathrm{cov}(X, Y) = 0$
  - yet, the two are not independent!
  - to see this, compute the conditional expectation $E[Y|X]$

Example: Zero Correlation, But Not Independent

- consider the following distribution:

$$f_{XY}(x,y) = \begin{cases} 3/4 & \text{if } x \in (-1,1) \text{ and } y \in (0, 1-x^2) \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

- show that $\mathrm{cov}(X,Y) = 0$
  - yet, the two are not independent!
  - to see this, compute the conditional expectation $E[Y|X]$

Example: Zero Correlation, But Not Independent

- consider the following distribution:

$$f_{XY}(x,y) = \begin{cases} 3/4 & \text{if } x \in (-1,1) \text{ and } y \in (0, 1-x^2) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

- show that $\text{cov}(X,Y) = 0$
  - yet, the two are not independent!
  - to see this, compute the conditional expectation $E[Y|X]$

Example: Zero Correlation, But Not Independent

- consider the following distribution:

$$f_{XY}(x,y) = \begin{cases} 3/4 & \text{if } x \in (-1,1) \text{ and } y \in (0, 1-x^2) \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

- show that $\text{cov}(X, Y) = 0$
  - yet, the two are not independent!
  - to see this, compute the conditional expectation $E[Y|X]$

Law of Iterated Expectations and Variance Decomposition

- law of iterated expectations:

$$E[E[Y|X]] = E[Y]$$

- variance decomposition:

$$V[Y] = V[E[Y|X]] + E[V[Y|X]]$$

Law of Iterated Expectations and Variance Decomposition

- law of iterated expectations:

$$E[E[Y|X]] = E[Y]$$

- variance decomposition:

$$V[Y] = V[E[Y|X]] + E[V[Y|X]]$$

Example: Censored Normal Distribution

Classical Statistical Paradigm

Modes of Convergence

Law(s) of Large Numbers

Central Limit Theorems

Desirable Properties of Estimators

References

Table of Contents