Stanley Wei

Math 170A (Killip, W24)

Study Guide

Contents

1 Probabilities

1.1 Probability Spaces

Definition (*Sample Space*). Given an experiment, we define its ***sample space*** Ω to be the set of all elementary outcomes.

Definition (σ-*Algebra*). A collection $\mathcal{F}$ of subsets $E \subseteq \Omega$ is called a σ***-algebra*** if the following hold:

1. $\emptyset \in \mathcal{F}$ and $\Omega \in \mathcal{F}$.
2. If $A \in \mathcal{F}$, $A^c \in \mathcal{F}$.
3. If $A_n \in \mathcal{F}$ for $n \in \mathbb{N}$, $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$. [*Countable unions*]

Definition (*Probability Measure*). Given a sample space Ω and σ-algebra $\mathcal{F}$, a ***probability [measure]*** is an assignment $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ subject to:

1. $\mathbb{P}(A) \geq 0 \;\forall\; A \in \mathcal{F}$.
2. $\mathbb{P}(\emptyset) = 0$; $\mathbb{P}(\Omega) = 1$.
3. If $A_n \in \mathcal{F}$ are disjoint for $n \in \mathbb{N}$, $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$. [$\sigma$-*additivity*]

1.2 Probabilities

DeMorgan's Laws:

1. $(A \cup B)^c = A^c \cap B^c$
2. $(A \cap B)^c = A^c \cup B^c$

Probability Properties:

Let $A, B \in \mathcal{F}$.

1. $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$
2. $0 \leq \mathbb{P}(A) \leq 1$
3. $(A \cap B) \in \mathcal{F}$.
4. $(A \setminus B) \in \mathcal{F}$.
5. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
6. If $A \subseteq B$, $\mathbb{P}(B) \geq \mathbb{P}(A)$
7. If $A_n \in \mathcal{F}$ for $1 \leq n \leq N$, then $\bigcup_{n=1}^{N} A_n \in \mathcal{F}$. [*Finite union*]

Theorem (**Continuity**).

1. Given events $A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n \subseteq \ldots$ for $n \in \mathbb{N}$, $\underline{\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)}$.
2. Given events $A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n \supseteq \ldots$ for $n \in \mathbb{N}$, $\underline{\mathbb{P}(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n)}$.

1.3 Conditional Probability

Definition (***Conditional Probability***)**.** Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and an event B with $\mathbb{P}(B) > 0$, then the ***conditional probability*** of A given B is:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

"**Multiplication Rule**":

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_n \mid A_1 \cap \ldots \cap A_{n-1}) \cdot \ldots \cdot \mathbb{P}(A_2 \mid A_1) \cdot \mathbb{P}(A_1)$$

1.4 Bayes' Rule

Theorem (***Partition Theorem***)**.** If $\{B_j\}$ form a partition of Ω with $\mathbb{P}(B_j) > 0 \ \forall\, j$, then $\forall\, A \in \mathcal{F}$:

$$\mathbb{P}(A) = \sum_j \mathbb{P}(A \cap B_j) = \sum_j \mathbb{P}(A \mid B_j) \cdot \mathbb{P}(B_j)$$

Definition (***Bayes' Rule***)**.** Given events A, B with $\mathbb{P}(A), \mathbb{P}(B) > 0$:

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(A \mid B) \cdot \mathbb{P}(B)}{\mathbb{P}(A)}$$

Definition (***Bayes' Rule on Partitions***)**.** Given a countable partition $\{B_j\}$ of Ω with $\mathbb{P}(B_j) > 0 \ \forall\, j$ and event A with $\mathbb{P}(A) > 0$:

$$\mathbb{P}(B_j \mid A) = \frac{\mathbb{P}(B_j \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \mid B_j)\,\mathbb{P}(B_j)}{\sum_k \mathbb{P}(A \mid B_k)\,\mathbb{P}(B_k)}$$

Given a partition $\{B_j\}$ of Ω and event A, distinguish between:

1. ***Likelihoods*** $\mathbb{P}(A \mid B_j)$
   - Do **not** comprise a probability measure
2. ***Posterior probabilities*** $\mathbb{P}(B_j \mid A)$
   - **Do** comprise a probability measure

1.5 Independence

Definition (***Independence***)**.** We say that:

(i) Two events A, B are ***[statistically] independent*** if:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$$

- Alternatively:

$$\mathbb{P}(A \mid B) = \mathbb{P}(A)$$

(ii) Events E_α for $\alpha \in A$ are ***independent*** if, for every finite $B \subseteq A$:

$$\mathbb{P}\left(\bigcap_{\alpha \in B} E_\alpha\right) = \prod_{\alpha \in B} \mathbb{P}(E_\alpha)$$

Definition (***Conditional Independence***)**.** Let $\mathbb{P}(C) > 0$; say that two events A, B are ***conditionally independent*** given C if:

$$\mathbb{P}(A \cap B \mid C) = \mathbb{P}(A \mid C)\,\mathbb{P}(B \mid C)$$

- Alternatively:

$$\mathbb{P}(A \mid B, C) = \mathbb{P}(A \mid C)$$

1.5.1 ($*$) Independence of Infinite Experiments

Given an infinite experiment, we can describe it via the infinite product of probability spaces.

Namely, given probability spaces $(\Omega_\alpha, \mathcal{F}_\alpha, \mathbb{P}_\alpha)$ for $\alpha \in A$, we can take the sample space $\prod_{\alpha \in A} \Omega_\alpha$.

Definition (***Cylinder Set***)**.** We define a ***cylinder set*** as follows: let $S \subseteq A$ be a finite set, and for each $\alpha \in S$ let $E_\alpha \in \mathcal{F}_\alpha$. Then the associated cylinder set is:

$$\left(\prod_{\alpha \in S} E_\alpha\right) \times \left(\prod_{\alpha \in S^c} \Omega_\alpha\right)$$

Theorem. Given probability spaces $(\Omega_\alpha, \mathcal{F}_\alpha, \mathbb{P}_\alpha)$ for $\alpha \in A$, there is a unique probability law on $(\prod_{\alpha \in A}, \otimes_{\alpha \in A} \mathcal{F}_\alpha, \mathbb{P})$ satisfying ($\forall$ finite $S \subseteq A$):

$$\mathbb{P}\left(\left(\prod_{\alpha \in S} E_\alpha\right) \times \left(\prod_{\alpha \in S} \Omega_\alpha\right)\right) = \prod_{\alpha \in S} \mathbb{P}_\alpha(E_\alpha)$$

2 Discrete Random Variables

Definition (***Discrete Random Variable***)**.** A ***discrete [real-valued] random variable*** on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a function

$$X : \Omega \to \mathbb{Z}$$

such that X is ***measurable***; i.e.: for each $n \in X(\Omega)$, $\{\omega \in \Omega : X(\omega) = n\} \in \mathcal{F}$.

Definition (***Probability Mass Function***)**.** Given discrete random variable X, its ***probability mass function*** [**PMF**] is the function $p_X : \mathbb{Z} \to [0, 1]$ defined by:

$$p_X(n) = \mathbb{P}\left(\{X = n\}\right)$$

Definition (***Indicator***)**.** Given an event $E \in \mathcal{F}$, its ***indicator random variable*** is the function 1_E defined by:

$$1_e(\omega) = \begin{cases} 1 & \omega \in E \\ 0 & \omega \notin E \end{cases}$$

2.1 Discrete Probability Distributions

Distribution	**Parameters**	**PMF**	$\mathbb{E}\left(X\right)$	$\mathrm{Var}\left(X\right)$
Bernoulli(p)	$p \in [0, 1]$	$p_X(k) = \begin{cases} p & k = 1 \\ (1-p) & k = 0 \\ 0 & \text{otherwise} \end{cases}$	p	$p(1-p)$
Binomial(n, p)	$n \in \mathbb{N} \cup \{0\}; p \in [0, 1]$	$p_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & 0 \leq k \leq n \\ 0 & \text{otherwise} \end{cases}$	np	$p(n - np)$
Geometric(p)	$p \in (0, 1]$	$p_X(k) = \begin{cases} (1-p)^{k-1} p & k = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson(λ)	$\lambda \in \mathbb{N} \cup \{0\}$	$p_X(k) = \begin{cases} \frac{\lambda^k}{k!} e^{-\lambda} & k \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases}$	λ	λ

Interpretations:

1. Bernoulli(p): One trial with probability of success p

2. Binomial(n, p): Number of successes across n independent Bernoulli(p) trials

3. Geometric(p): Number of independent Bernoulli(p) trials to reach 1st success

4. Poisson(λ): Expected number of occurrences of some event across some interval

2.2 Expected Value

Definition (***Expected Value***)**.** Given a discrete r.v. X, we define its ***expected value*** (alt: mean, average, expectation) to be:

$$\mathbb{E}(X) = \sum_k k \cdot p_X(k),$$

given that the sum converges absolutely ($\mathbb{E}(X) = \sum_k |\, k \cdot p_X(k) \,| < \infty$).

Expected Value & Composition:

$$\mathbb{E}(g(X)) = \sum_k g(k)\mathbb{P}(X = k)$$

Theorem (***Jensen's inequality***)**.** If $g : \mathbb{R} \to \mathbb{R}$ is convex, then:

$$\mathbb{E}(g(X)) \geq g(\mathbb{E}(X)$$

2.3 Variance & Standard Deviation

Definition (*Variance & Standard Deviation*). Given random variable X:

(i) Call $\mathbb{E}(X^m)$ the m^{th} moment of X.

(ii) Define the ***variance*** of X as $\mathrm{Var}(X) = \mathbb{E}((X-\mu)^2)$, where $\mu = \mathbb{E}(X)$.

(iii) Define the ***standard deviation*** of X as $\sigma(X) = \sqrt{\mathrm{Var}(X)}$.

Properties:

(i) $\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$

(ii) $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$

2.4 Conditional Expectation

Definition (*Conditional Expectation*). If X a random variable and A an event with $\mathbb{P}(A) > 0$, define the ***conditional PMF*** of X given A:

$$p_{X\,|\,A} = \mathbb{P}(X = k \,|\, A)$$

Similarly, we define the ***conditional expectation*** of X given A:

$$\mathbb{E}(X \,|\, A) = \sum_k k \cdot p_{X\,|\,A}(k),$$

provided the sum converges absolutely.

Definition (*Partition Theorem*). Given r.v. X with $\mathbb{E}(|X|) < \infty$ and partition $\{B_j\}$ of Ω with each $\mathbb{P}(B_j) \geq 0$, we have:

$$\mathbb{E}(X) = \sum_j \mathbb{E}(X \,|\, B_j)\,\mathbb{P}(B_j)\,.$$

In particular, the expectations $\mathbb{E}(X \,|\, B_j)$ exist.

3 Multiple Discrete Random Variables

Definition (***Joint PMF***)**.** Given two random variables X, Y, the ***joint PMF of X and Y*** is the function $p_{X,Y}$ defined by:

$$p_{X,Y}(k,l) = \mathbb{P}\left(X = k \;\&\; Y = l\right)$$

Definition (***Independence***)**.** We say random variables X, Y are ***independent*** if, for <u>any</u> choice of k, l, the events $\{X = k\}$ and $\{Y = l\}$ are independent, i.e.:

$$\mathbb{P}\left(X = k \;\&\; Y = l\right) = \mathbb{P}\left(X = k\right)\mathbb{P}\left(Y = l\right)$$

Independence of Multiple Variables

- We say random variables $X_1, X_2, \ldots$ are ***independent*** if, for any choice of k_i's for $i \in \mathbb{N}$, the events $\{X_i = k_i\}$ are independent.
- If two random variables X, Y are independent and have expectation, then the variable XY has finite expectation $\mathbb{E}\left(XY\right) = \mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)$.

Covariance

Definition (***Covariance***)**.** Given two r.v.s X and Y, the ***covariance*** of X and Y is:

$$\operatorname{Cov}\left(X, Y\right) = \mathbb{E}\left(\left(X - \mathbb{E}\left(X\right)\right)\left(Y - \mathbb{E}\left(Y\right)\right)\right)$$

In particular: $\operatorname{Cov}\left(X, Y\right) = \mathbb{E}\left(XY\right) - \mathbb{E}\left(X\right)\mathbb{E}\left(Y\right)$

4 Continuous Random Variables

Definition (***Real-Valued Random Variable***)**.** A real-valued random variable on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a function $X : \Omega \to \mathbb{R}$ that is measurable, i.e.: for all $x \in \mathbb{R}$, $\{X \leq x\} \in \mathcal{F}$.

4.1 Cumulative Distribution Functions

Definition (***Cumulative Distribution Function***)**.** Given real-valued random variable X, its ***cumulative distribution function*** [**CDF**] is the function $F_x : \mathbb{R} \to [0, 1]$ defined by:

$$F_X(x) = \mathbb{P}\left(X \leq x\right)$$

CDFs

- CDFs are always strictly non-decreasing
- CDFs may be discrete, continuous, or neither discrete nor continuous
- CDFs are "cadlag": continuous from the left, limit from the right:
  - $\lim_{y \uparrow x} F_X(y) = \mathbb{P}\left(X \leq x\right)$
  - $\lim_{y \downarrow x} F_X(y) = F_X(x)$

4.2 Continuous Random Variables

Definition (***Continuous Random Variables***)**.** A random variable $X : \Omega \to \mathbb{R}$ is called ***continuous*** if $\exists$ integrable function $f_X : \mathbb{R} \to \mathbb{R}$ such that:

$$F_X(x) = \int_{-\infty}^{x} f_X(x')dx'$$

The function f_X is called the ***probability density function*** [**PDF**] of X.

Definition (***Absolute Continuity***)**.** A function F is called ***absolutely continuous*** if, for every $\epsilon > 0$, $\exists\, \delta > 0$ s.t. for any set of intervals $((a_i, b_i))_i$ with $|b_1 - a_1| + \ldots + |b_n - a_n| < \delta$, then:

$$\sum_{i=1}^{n} |F(b_i) - F(a_i)| < \epsilon$$

Continuous Random Variables

- It is true that for any continuous R.V. X, its CDF F_X is continuous; however, it is not true that every continuous CDF F_X admits a probability density
  - Only absolutely continuous CDFs admit a probability density
- If $X : \Omega \to \mathbb{R}$ is a continuous r.v., then $\forall\, a \leq b$:

$$\mathbb{P}(a < X < b) = \mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)dx$$

(∗) Corollary: $\mathbb{P}(X = a) = 0 \;\forall\, a \in \mathbb{R}$

4.3 Continuous Probability Distributions

Definition (*Gamma Function*). For $z > 0$, the ***gamma function*** $\Gamma(z)$ is defined by:

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$$

(∗) *Note*: $\underline{\Gamma(s+1) = s\Gamma(s)}$; for $n \in \mathbb{N}$, $\Gamma(n+1) = n!$

Distribution	**Parameters**	**PDF**	$\mathbb{E}(X)$	$\mathrm{Var}(X)$
Uniform($[a, b]$)	$a \leq b$	$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$	$\frac{a+b}{2}$	$\frac{1}{12}(b-a)^2$
$\mathrm{N}(\mu, \sigma^2)$	μ; $\sigma \geq 0$	$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Exponential(λ)	$\lambda > 0$	$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$\Gamma(k, \lambda)$	$k, \lambda > 0$	$f_x(x) = \begin{cases} \frac{1}{\Gamma(k)} \lambda^k x^{k-1} e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$\frac{k}{\lambda}$	$\frac{k}{\lambda^2}$
Cauchy	None	$f_X(x) = \frac{1}{\pi(1+x^2}$	None	None

Additional Distributions:

- $W \sim \text{LogNormal}(\mu, \sigma^2)$ if $\ln(W) \sim \text{N}(\mu, \sigma^2)$
- If $X_1, \dots, X_n \sim \text{Exp}(\lambda)$ are independent, then $\sum_{i=1}^n X_i \sim \text{Erlang}(n, \lambda) = \Gamma(n, \lambda)$
- If $X_1, \dots, X_n \sim \text{N}(0, 1)$ independent, then $\sum_{i=1}^n X_i \sim \chi_n^2 = \Gamma(\frac{n}{2}, \frac{1}{2})$

4.4 Expectation for Continuous Random Variables

Definition (***Expectation***)**.** Given a continuous random variable X, we define its ***expectation*** to be:

$$\mathbb{E}\left(X\right) = \int_{-\infty}^{\infty} x \cdot f_X(x) dx$$

provided the integral is absolutely convergent.

Expectation for Continuous Random Variables

Given a continuous function g and continuous random variable X:

- $g \circ X$ is a random variable, i.e. $g \circ X$ is measurable
- $g \circ X$ is not necessarily a continuous random variable
- $g \circ X$ has expectation (if it exists) given by:

$$\mathbb{E}\left(g(X)\right) = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

4.5 (∗) Generalized Quantile Functions

Definition (***Generalized Quantile Function***)**.** Given a CDF F, the ***generalized quantile function*** associated to F is the function $Q : (0, 1) \to \mathbb{R}$ defined by:

$$Q(p) = \inf\{x : F(x) \geq p\}$$

Generalized Quantile Functions

- If F is invertible, then $Q = F^{-1}$
- $Q(p) \leq x \Leftrightarrow F(x) \geq p$

5 Multiple Continuous Random Variables

5.1 Joint CDF & Independence

Definition (***Joint CDF***)**.** Given 2 random variables $X, Y : \Omega \to \mathbb{R}$, the ***joint CDF*** of X and Y is the function $\underline{F_{X,Y} : \mathbb{R}^2 \to [0,1]}$ defined by:

$$F_{X,Y}(x,y) = \mathbb{P}\left(X \leq x \,\&\, Y \leq y\right)$$
$$\Updownarrow$$
$$F_{X,Y}(x,y) = \mathbb{P}\left(\{\omega : X(\omega) \leq x\} \cap \{\omega : Y(\omega) \leq y\}\right)$$

Definition (***Independence***)**.** We say that two random variables X, Y are ***independent*** if, $\forall\, x, y \in \mathbb{R}$:

$$\mathbb{P}\left(X \leq x \,\&\, Y \leq y\right) = \mathbb{P}\left(X \leq x\right)\mathbb{P}\left(Y \leq y\right)$$
$$\Updownarrow$$
$$F_{X,Y}(x,y) = F_X(x)F_Y(y)$$

5.2 Joint Continuity

Definition (***Joint Continuity***)**.** We say that two random variables X, Y are ***jointly continuous*** if there is an $\underline{\text{integrable function } f_{X,Y} : \mathbb{R}^2 \to [0,1]}$:

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x', y') dy' dx'$$

The function $f_{X,Y}$, if it exists, is called the ***joint PDF*** of X and Y.

Joint Continuity

- If X, Y are jointly continuous, then they are individually continuous.
- If X, Y are jointly continuous, then we can find the ***marginal distribution*** of X (and likewise for Y) as:

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

- If two variables X, Y are individiually continuous and independent, then they are jointly continuous with joint PDF $f_{X,Y}(x,y) = f_X(x)f_Y(y)$
- If X, Y are jointly continuous, then $\mathbb{E}(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy$.

5.3 Conditional Density Functions

Definition (***Conditional CDF***)**.** Given a continuous random variable X and an event A with $\mathbb{P}(A) > 0$, the ***conditional CDF*** of X given A is:

$$\mathbb{P}(X \leq x \,|\, A) = \frac{\mathbb{P}(X \leq x \,\&\, A)}{\mathbb{P}(A)}$$

Definition (***Conditional PDF***)**.** Given continuous random variables X, Y, the ***conditional PDF*** of Y given X is:

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

Definition (***Conditional Expectation***)**.** The ***conditional expectation*** of Y given X is:

$$\mathbb{E}(Y \,|\, X = x) = \int_{-\infty}^{\infty} y f_Y \,|\, X(y \,|\, x)dy$$

Theorem. Given continuous random variables X, Y:

1. **Law of Total Probability**: $f_Y(y) = \int_X f_{Y\,|\,X}(y \,|\, x) f_X(x)dx$
2. **Law of Total Expectation**: $\mathbb{E}(Y) = \int_X \mathbb{E}(Y \,|\, X = x) f_X(x)dx$

Conditional Density & Expectation

- $\mathbb{E}(g(Y) \,|\, X = x) = \int_X g(y) f_{Y\,|\,X}(y \,|\, x)dx$
- $\mathbb{E}(g(Y)h(X)) = \int_X \mathbb{E}(g(Y) \,|\, X = x) h(x) f_X(x)dx$
- Two jointly continuous random variables X, Y are independent iff $f_{Y\,|\,X}(y \,|\, x) = f_Y(y)$ for all but a negligible set $x \in E$

5.4 Multivariate Normal Distribution

Definition (***Positive-Definite***)**.** A real matrix Σ is called ***positive-definite*** if it is square, symmetric, and has $\forall\, \vec{\xi} \in \mathbb{R}^n \setminus \{\vec{0}\}$:

$$\vec{\xi} \bullet \Sigma \vec{\xi} > 0$$

Theorem. For any positive-definite matrix Σ, $\exists$ invertible positive-definite matrix B s.t.:

$$\Sigma = BB^T$$

Definition (***Multivariate Normal Distribution***)**.** We say that random variables $X_1, \ldots, X_n$ are ***jointly normal*** if there exist:

1. A positive-definite matrix Σ (called the ***covariance matrix***) and
2. A vector $\vec{\mu} \in \mathbb{R}^n$ (called the ***mean***)

such that $\vec{X} = \langle \vec{X}_1, \ldots, \vec{X}_n \rangle$ has PDF given by:

$$f_{\vec{X}}(\vec{x}) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu}) \bullet \Sigma^{-1}(\vec{x} - \vec{\mu})\right)$$

($*$) *Notation*: Say that: $\int f(\vec{x})d\vec{x} = \int \ldots \int f(\vec{x})dx_n \ldots dx_1$, where $\vec{x} = \langle x_1, \ldots, x_n \rangle$

Theorem. Given $\vec{X}, \Sigma, \mu$ as above:

1. $\int f_{\vec{X}}(\vec{x})d\vec{x} = 1$
2. Mean: $\int \vec{x} f_{\vec{X}}(\vec{x})d\vec{x} = \vec{\mu}$
3. Covariance: $\operatorname{Cov}(X_i, X_j) = \int (\vec{x}_i - \vec{\mu}_i)(\vec{x}_j - \vec{\mu}_j) f_{\vec{X}}(\vec{x})d\vec{x} = \Sigma_{ij}$