VE401 Probabilistic Methods in Eng.
RC 3

CHEN Xiwen

UM-SJTU Joint Institute

March 19, 2020

Table of contents

More About Normal Distribution
- Normal Distribution
- Applications of Normal Distribution

Multivariate Random Variables
- Discrete Multivariate Random Variables
- Continuous Multivariate Random Variables
- Expectation and Variance
- The Hypergeometric Distribution

Exercises

More About Normal Distribution

Normal Distribution

Applications of Normal Distribution

Multivariate Random Variables

Discrete Multivariate Random Variables

Continuous Multivariate Random Variables

Expectation and Variance

The Hypergeometric Distribution

Exercises

Normal Distribution

Definition. A continuous random variable (X, f_{μ,σ^2}) has the ***normal distribution*** with mean $\mu \in \mathbb{R}$ and variance $\sigma^2, \sigma > 0$ if the probability density function is given by

$$f_{\mu,\sigma^2} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \qquad x \in \mathbb{R}.$$

Normal Distribution

Mean, variance and M.G.F.

- Mean.

$$\mathrm{E}[X] = \mu.$$

- Variance.

$$\mathrm{Var}[X] = \sigma^2.$$

- M.G.F.

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right).$$

Normal Distribution

Verifying M.G.F.

$$
\begin{aligned}
m_X(t) = \mathrm{E}\left[e^{tX}\right] &= \int_{-\infty}^{\infty} \frac{e^{tx}}{\sqrt{2\pi}\sigma} e^{-((x-\mu)/\sigma)^2/2} \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{\mu t+\sigma^2 t^2/2} \cdot e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}} \mathrm{d}x \\
&= e^{\mu t+\sigma^2 t^2/2} \underbrace{\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-\frac{(x-(\mu+\sigma^2 t))^2}{2\sigma^2}} \mathrm{d}x}_{=1} \\
&= e^{\mu t+\sigma^2 t^2/2}.
\end{aligned}
$$

Normal Distribution

Some takeaway from this proof.

- To verify that

$$I := \int_{-\infty}^{\infty} e^{-\frac{(x-b)^2}{a^2}} \mathrm{d}x = a\sqrt{\pi},$$

we use

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{b^2}} \mathrm{d}x\right)^2 = \int_{-\infty}^{\infty} e^{-\frac{(x-a)^2}{b^2}} \cdot e^{-\frac{(y-a)^2}{b^2}} \mathrm{d}x\mathrm{d}y.$$

Using parametrization $x = ar\cos\theta + b, y = ar\sin\theta + b$, we have

$$\begin{aligned} I^2 &= \int_0^{\infty} \int_0^{2\pi} e^{-r^2} \cdot a^2 r \mathrm{d}\theta \mathrm{d}r \\ &= a^2\pi \int_0^{\infty} 2re^{-r^2} \mathrm{d}r = -a^2\pi e^{-r^2}\Big|_0^{\infty} = a^2\pi. \end{aligned}$$

Normal Distribution

Some takeaway from this proof.

- Useful results from normalizing constant of distributions.
  (i). Normal.

$$\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x = \sqrt{2\pi}\sigma.$$

  (ii). Gamma.

$$\int_{0}^{\infty} x^{\alpha-1} e^{-\beta x} \mathrm{d}x = \frac{\Gamma(\alpha)}{\beta^\alpha}.$$

Transformation of Random Variables

- Discrete random variables. Let X be a discrete random variable with probability density function f_X, the the probability density function f_Y for $Y = \varphi(X)$ is given by

$$f_Y(y) = \sum_{x \in \varphi^{-1}(y)} f_X(x), \qquad \text{for } y \in \text{ran } \varphi,$$

and 0 otherwise.

Example. Let X be a uniform random variable on $\{-n, -n+1, \ldots, n-1, n\}$. Then $Y = |X|$ has probability density function

$$f_Y(y) = \begin{cases} \dfrac{1}{2n+1} & x = 0, \\ \dfrac{2}{2n+1} & x \neq 0. \end{cases}$$

Transformation of Random Variables

- Continuous random variables. Let X be a continuous random variable with density f_X. Let $Y = \varphi \circ X$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is strictly monotonic and differentiable. The density for Y is then given by

  $$f_Y(y) = f_X(\varphi^{-1}(y)) \cdot \left| \frac{\mathrm{d}\varphi^{-1}(y)}{\mathrm{d}y} \right|, \qquad \text{for } y \in \operatorname{ran} \varphi$$

  and

  $$f_Y(y) = 0, \qquad \text{for } y \notin \operatorname{ran} \varphi.$$

Standardizing Normal Distribution

Suppose $X \sim \text{Normal}(\mu, \sigma^2)$. Then

$$Z = \frac{X - \mu}{\sigma} \sim \text{Normal}(0, 1),$$

where the normal distribution with mean μ and variance σ^2 is the ***standard normal distribution***. Furthermore, the cumulative distribution function of X is given by

$$F(x) = \Phi\left(\frac{x - \mu}{\sigma}\right), \quad F^{-1}(p) = \mu + \sigma\Phi^{-1}(p),$$

where Φ is the cumulative distribution function for the standard normal distribution function.

More About Normal Distribution
Normal Distribution
Applications of Normal Distribution

Multivariate Random Variables
Discrete Multivariate Random Variables
Continuous Multivariate Random Variables
Expectation and Variance
The Hypergeometric Distribution

Exercises

Common Applications of Normal Distribution

Suppose a random variable X follows normal distribution $N(\mu, \sigma)$, where μ and σ are known. At current stage, applications usually include the following.

1. Given some value x_0, find the probability of $P[X \leq x_0]$ or $P[X \geq x_0]$.
   (a). Standardize X as $Z = (X - \mu)/\sigma$, find z_0.
   (b). Find $P[X \leq x_0] = P[Z \leq z_0], P[X \geq x_0] = 1 - P[Z \geq z_0]$.
2. Given some probability p, find the corresponding x_0 such that $P[X \leq x_0] = p$ or $P[X \geq x_0] = p$.
   (a). Find z_0 from table such that $P[Z \leq z_0] = p$ or $P[Z \leq z_0] = 1 - p$.
   (b). Calculate $x_0 = \sigma z_0 + \mu$.
3. "Three-sigma" rule.

$$P[-3\sigma < X - \mu < 2\sigma] = 0.997.$$

Normal Approximation of Binomial Distribution

- Theorem. Suppose $X \sim \text{Binomial}(n, p)$, then

$$\lim_{n\to\infty} P\left[a < \frac{X - np}{\sqrt{np(1-p)}} \leq b\right] = \frac{1}{2\pi}\int_a^b e^{-x^2/2}\mathrm{d}x.$$

- Suppose $X \sim \text{Binomial}(n, p)$, then for $y = 0, \ldots, n$,

$$P[X \leq y] = \sum_{x=0}^{y} \binom{n}{x} p^x (1-p)^{n-x} \approx \Phi\left(\frac{y + 1/2 - np}{\sqrt{np(1-p)}}\right),$$

where we require that

$$np > 5 \quad \text{if } p \leq \frac{1}{2} \qquad \text{or} \qquad n(1-p) > 5 \quad \text{if } p > \frac{1}{2}.$$

Lyapunov's Central Limit Theorem

Theorem. Let (X_i) be a sequence of independent, but not necessarily identical random variables whose moments exist and satisfy a certain technical condition. Let

$$Y_n = X_1 + \cdots + X_n.$$

Then for any $z \in \mathbb{R}$,

$$P\left[\frac{Y_n - \mathrm{E}[Y_n]}{\sqrt{\mathrm{Var}[Y_n]}} \leq z\right] \xrightarrow{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\mathrm{d}x.$$

More About Normal Distribution
Normal Distribution
Applications of Normal Distribution

Multivariate Random Variables
Discrete Multivariate Random Variables
Continuous Multivariate Random Variables
Expectation and Variance
The Hypergeometric Distribution

Exercises

Discrete Multivariate Random Variables

Definition. Let S be a sample space and Ω a countable subset of $\mathbb{R}^n$. A ***discrete multivariate random variable*** is a map

$$\mathbf{X} : S \to \Omega$$

together with a function $f_{\mathbf{X}} : \Omega \to \mathbb{R}$ with the properties that

(i). $f_{\mathbf{X}}(x) \geq 0$ for all $x = (x_1, \ldots, x_n) \in \Omega$ and

(ii). $\displaystyle\sum_{x \in \Omega} f_{\mathbf{X}}(x) = 1$,

where $f_{\mathbf{X}}$ is the ***joint density function*** of the random variable $\mathbf{X}$.

Discrete Multivariate Random Variables

Definition.

- ***Marginal density*** f_{X_k} for $X_k, k = 1, \ldots, n$:

$$f_{X_k}(x_k) = \sum_{x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n} f_{\mathbf{X}}(x_1, \ldots, x_n).$$

- ***Independent*** multivariate random variables:

$$f_{\mathbf{X}}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

- ***Conditional density*** of X_1 conditioned on X_2:

$$f_{X_1|X_2}(x_1) := \frac{f_{X_1X_2}(x_1, x_2)}{f_{X_2}(x_2)} \qquad \text{whenever } f_{X_2}(x_2) > 0.$$

More About Normal Distribution
Normal Distribution
Applications of Normal Distribution

Multivariate Random Variables
Discrete Multivariate Random Variables
Continuous Multivariate Random Variables
Expectation and Variance
The Hypergeometric Distribution

Exercises

Continuous Multivariate Random Variables

Definition. Let S be a sample space. A ***continuous multivariate random variable*** is a map

$$\mathbf{X} : S \to \mathbb{R}^n$$

together with a function $f_{\mathbf{X}} : \mathbb{R}^n \to \mathbb{R}$ with the properties that

(i). $f_{\mathbf{X}}(x) \geq 0$ for all $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and

(ii). $\displaystyle\int_{\mathbb{R}^n} f_{\mathbf{X}}(x) = 1,$

where $f_{\mathbf{X}}$ is the ***joint density function*** of the random variable $\mathbf{X}$.

Continuous Multivariate Random Variables

Definition.

- ***Marginal density*** f_{X_k} for $X_k, k = 1, \dots, n$:

$$f_{X_k}(x_k) = \int_{\mathbb{R}^{n-1}} f_{\mathbf{X}}(x_1, \dots, x_n) \mathrm{d}x_1 \dots \mathrm{d}x_{k-1} \mathrm{x}_{\mathrm{k}+1} \dots \mathrm{d}x_n.$$

- ***Independent*** multivariate random variables:

$$f_{\mathbf{X}}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

- ***Conditional density*** of X_1 conditioned on X_2:

$$f_{X_1|X_2}(x_1) := \frac{f_{X_1 X_2}(x_1, x_2)}{f_{X_2}(x_2)} \qquad \text{whenever } f_{X_2}(x_2) > 0.$$

More About Normal Distribution
Normal Distribution
Applications of Normal Distribution

Multivariate Random Variables
Discrete Multivariate Random Variables
Continuous Multivariate Random Variables
Expectation and Variance
The Hypergeometric Distribution

Exercises

Expectation

- Discrete.

$$\mathrm{E}[X_k] = \sum_{x_k} x_k f_{X_k}(x_k) = \sum_{x\in\Omega} x_k f_{\mathbf{X}}(x),$$

 and for continuous function $\varphi : \mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{E}[\varphi \circ \mathbf{X}] = \sum_{x\in\Omega} \varphi(x) f_{\mathbf{X}}(x).$$

- Continuous.

$$\mathrm{E}[X_k] = \int_{\mathbb{R}} x_k f_{X_k}(x_k)\mathrm{d}x_k = \int_{\mathbb{R}^n} x_k f_{\mathbf{X}}(x)\mathrm{d}x,$$

 and for continuous function $\varphi : \mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{E}[\varphi \circ \mathbf{X}] = \int_{\mathbb{R}^n} \varphi(x) f_{\mathbf{X}}(x)\mathrm{d}x.$$

Covariance and Covariance Matrix

Definition. For a multivariate random variable $\mathbf{X}$, the ***covariance matrix*** is given by

$$\mathrm{Var}[\mathbf{X}] = \begin{pmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] & \cdots & \mathrm{Cov}[X_1, X_n] \\ \mathrm{Cov}[X_1, X_2] & \mathrm{Var}[X_2] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{Cov}[X_{n-1}, X_n] \\ \mathrm{Cov}[X_1, X_n] & \cdots & \mathrm{Cov}[X_{n-1}, X_n] & \mathrm{Var}[X_n] \end{pmatrix},$$

where the ***covariance*** of (X_i, X_j) is given by

$$\mathrm{Cov}[X_i, X_j] = \mathrm{E}[(X_i - \mu_{X_i})(X_j - \mu_{X_j})],$$

and

$$\mathrm{Var}[\mathbf{CX}] = \mathbf{C}\mathrm{Var}[\mathbf{X}]\mathbf{C}^T, \qquad \mathbf{C} \in \mathrm{Mat}(n \times n; \mathbb{R}).$$

Covariance and Independence

Let $X, X_1, \ldots, X_n$ and Y be random variables.

- X and Y are independent $\Rightarrow \text{Cov}[X, Y] = 0$, while the converse is not true.
- $\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X, Y]$, and more generally,

$$\text{Var}[X_1 + \cdots + X_n] = \text{Var}[X_1] + \cdots + \text{Var}[X_n] + \\ + 2\sum_{i<j} \text{Cov}[X_i, X_j],$$

if $\text{Var}[X_i] < \infty$ for $i = 1, \ldots, n$.

Covariance and Independence

Example. Suppose the random variable X can take only three values -1, 0, and 1, and each of these values has the same probability. Also, let random variable Y satisfy $Y = X^2$. Then X and Y are apparently dependent, while

$$\mathrm{E}[XY] = \mathrm{E}[X^3] = \mathrm{E}[X] = 0,$$

and thus

$$\mathrm{Cov}[X, Y] = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] = 0.$$

Pearson Correlation Coefficient

Definition. The ***Pearson coefficient of correlation*** of random variables X and Y is given by

$$\rho_{XY} := \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\mathrm{Var}[Y]}}.$$

Note. Instead of independence, the correlation coefficient actually measures the the extent to which X and Y are <u>linearly</u> dependent, which is not the only way of being dependent.

Properties.

(i). $-1 \leq \rho_{XY} \leq 1$,

(ii). $|\rho_{XY}| = 1$ iff there exist $\beta_0, \beta_1 \in \mathbb{R}$ such that

$$Y = \beta_0 + \beta_1 X.$$

The Fisher Transformation

Definition. Let $\tilde{X}$ and $\tilde{Y}$ be standardized random variables of X and Y, then the ***Fisher transformation*** of ρ_{XY} is given by

$$\ln\left(\sqrt{\frac{\mathrm{Var}[\tilde{X}+\tilde{Y}]}{\mathrm{Var}[\tilde{X}-\tilde{Y}]}}\right) = \frac{1}{2}\ln\left(\frac{1+\rho_{XY}}{1-\rho_{XY}}\right) = \mathrm{Arctanh}(\rho_{XY}) \in \mathbb{R}.$$

We say that X and Y are

- ***positively correlated*** if $\rho_{XY} > 0$, and
- ***negatively correlated*** if $\rho_{XY} < 0$.

More About Normal Distribution
Normal Distribution
Applications of Normal Distribution

Multivariate Random Variables
Discrete Multivariate Random Variables
Continuous Multivariate Random Variables
Expectation and Variance
The Hypergeometric Distribution

Exercises

The Hypergeometirc Distribution

Definition. A random variable (X, f_X) with parameters $N, n, r \in \mathbb{N}\backslash\{0\}$ where $r, n \leq N$ and $n < \min\{r, N-r\}$ has a ***hypergeometric distribution*** if the density function is given by

$$f_X(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}.$$

Interpretation.

- $f_X(x)$ is the probability of getting x balls in drawing n balls from a box containing N balls, where r of them are red.
- $f_X(x)$ is the probability of x successes in n identical but ***not*** independent Bernoulli trials, each with probability of success $\frac{r}{N}$.

The Hypergeometirc Distribution

- Expectation.

$$\mathrm{E}[X] = \mathrm{E}[X_1 + \cdots + X_n] = n\frac{r}{N}.$$

- Variance.

$$\begin{aligned}\mathrm{Var}[X] &= \mathrm{Var}[X_1 + \cdots + X_n] \\ &= \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n] + 2\sum_{i<j}\mathrm{Cov}[X_i, X_j] \\ &= n\frac{r}{N}\frac{N-r}{N}\frac{N-n}{N-1}.\end{aligned}$$

The binomial distribution may be used to approximate the hypergeometric distribution if n/N is small.

Thanks for your attention!