VE401 Recitation Class Note

Midterm Part1

Chen Siyi
siyi.chen_chicy@sjtu.edu.cn

Contents

1 Self-check

1. Elementary Probability
2. Conditional Probability
3. Discrete Random Variables
4. Continuous Random Variables
5. Expectation, Variance and Moments
6. Meanings of and Relationship Between Distributions
7. Multivariate Random Variables
8. Transformation of Random Variables
9. Reliability
10. Samples and Data Visualization
11. Parameter Estimation
12. Interval Estimation

2 Elementary Probability

2.1 Cardano's Principle

Let A be a random outcome of an experiment that may proceed in various ways. Assume each of these ways is **equally likely.** Then the probability P[A] of the outcome A is:

$$P[A] = \frac{number\, of\, ways\, leading\, to\, outcome\, A}{total\, number\, of\, ways\, (the\, experiment\, can\, proceed)}$$

2.2 Counting Ways for Events

Suppose a set A of n objects is given. From A:

1. **choose k ordered objects:** $\frac{n!}{(n-k)!}$
2. **choose k unordered objects:** $\frac{n!}{k!(n-k)!}$
3. **partitioning A into k disjoint subsets** $A_1, ..., A_k$ **having** $n_1, ..., n_k$ **elements:** $\frac{n!}{n_1!...n_k!}$

Understand: Order the total n objects, use clapboard to separate, consider repetition.

2.3 Axiomatic Approach

$\sigma-$Field

S is a non-empty set. A $\sigma-$field $\mathcal{F}$ on S is a set of subsets of S such that:

(i) $\emptyset \in \mathscr{F}$

(ii) if $A \in \mathscr{F}$, then $S \setminus A \in \mathscr{F}$

(iii) if $A_1, A_2, ... \in \mathscr{F}$ is a finite or countable sequence of subsets, then the union $\bigcup_k A_k \in \mathscr{F}$.

Probability Measures and Spaces

Let S be a sample space and $\mathscr{F}$ a $\sigma-$field on S. A **probability measure** (or **probability function** or just **probability**) on S is a **function** such that:

$$P : \mathscr{F} \to [0, 1], \quad A \mapsto P[A]$$

(i) P[S]=1

(ii) For any set of events $A_k \subset \mathscr{F}$ such that $A_j \cap A_k = \emptyset$ for $j \neq k$,

$$P\left[\bigcup_k A_k\right] = \sum_k P[A_k]$$

Then $(S, \mathscr{F}, P)$ is called a **probability space**.

Almost Surely

A $\in \sigma-$field, and P[A] = 1.
This does not means A = S.

Properties of Probability

$$P[S] = 1$$
$$P[\emptyset] = 0$$
$$P[S \setminus A] = 1 - P[A]$$
$$P[A_1 \cup A_2] = P[A_1] + P[A_2] - P[A_1 \cap A_2]$$

3 Conditional Probability

3.1 Definition

Given event A occurs, the probability for B to occur:

$$P[B|A] := \frac{P[A \cap B]}{P[A]}$$

3.2 Independence

Event A and event B are independent if:

$$P[A \cap B] = P[A] \cdot P[B]$$

And then other properties:

$$P[B|A] = P[B],\ if\ P[A] \neq 0$$

$$P[A|B] = P[A],\ if\ P[B] \neq 0$$

3.3 Total Probability

$$P[B] = \sum_{k=1}^{n} P\left[B|A_k\right] \cdot P\left[A_k\right]$$

3.4 Bayes's Theorem

$A_1, ..., A_n \subset S$ and pairwise mutually exclusive;
$\bigcup_n A_n = S$;
$B \subset S$ and $P[B] \neq 0$. Then:

$$P\left[A_k|B\right] = \frac{P\left[B \cap A_k\right]}{P[B]} = \frac{P\left[B|A_k\right] \cdot P\left[A_k\right]}{\sum_{j=1}^{n} P\left[B|A_j\right] \cdot P\left[A_j\right]}$$

4 Discrete Random Variables

4.1 Definition

Let S be a sample space and Ω a countable subset of R. A discrete random variable is a map

$$X : S \to \Omega$$

together with a function

$$f_X : \Omega \to \mathbb{R}$$

where

(i) $f_X(x) \geq 0$ for all $x \in \Omega$

(ii) $\sum_{x\in\Omega} f_X(x) = 1$

We often say that a random variable is given by the pair (X , f_X).

4.2 General Properties

1. $f_X = P[X = x]$: probability density function(PDF).
2. $F_X(x) = \sum_{y\leq x} f_X(y)$: cumulative distribution function(CDF).
3. $\mathrm{E}[X] := \sum_{x\in\Omega} x \cdot f_X(x)$: expectation.
4. $\mathrm{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = E[X^2] - E[X]^2$: Variance.
5. $m_X(t) := \sum_{k=0}^{\infty} \frac{E[X^k]}{k!} t^k = E[e^{tX}]$: moment generating function (MGF).

4.3 Expectation

Definition

$$\mathrm{E}[X] := \sum_{x\in\Omega} x \cdot f_X(x)$$

Exists only if E[X] converges.

Properties

1. $\mathrm{E}[\varphi \circ X] = \sum_{x\in\Omega} \varphi(x) \cdot f_X(x)$

2. $c \in \mathbb{R}$, then E[c]=c, E[cX]=cE[X]
3. X, Y both be random variables, then E[X+Y]=E[X]+E[Y]

Comments:

1. Describe the location of the average value
2. Different from "median"(or modes)

4.4 Variance

Definition

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

Standard deviation

$$\sigma_X = \sqrt{Var[X]}$$

Properties

1. $c \in \mathbb{R}$, then Var[c]=0, Var[cX]=c^2Var[X]
2. $Var[X+Y] = Var[X] + Var[Y] + 2Cov[X,Y]$
   $Cov[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] = E[XY] - E[X]E[Y]$

4.5 Moment Generating Functions

Moments

n_{th} (ordinary) moments of X: n=1,2,3,...

$$E[X^n]$$

n_{th} central moments of X: n=3,4,5,...

$$E\left[(\frac{X-\mu}{\sigma})^n\right]$$

MGF

$$m_X(t) := \sum_{k=0}^{\infty} \frac{E\left[X^k\right]}{k!} t^k = E[e^{tX}]$$

Radius of convergence

$$\varepsilon > 0$$

Application

$$\mathrm{E}\left[X^k\right] = \left.\frac{d^k m_X(t)}{dt^k}\right|_{t=0}$$

4.6 Summary of discrete distributions

distribution pmf	mean	variance	mgf
Poisson $e^{-\lambda}\frac{\lambda^x}{x!}$	λ	λ	$e^{-\lambda+\lambda e^t}$
Binomial $\binom{n}{x}\pi^x(1-\pi)^{n-x}$	$n\pi$	$n\pi(1-\pi)$	$(\pi e^t + 1 - \pi)^n$
Geometric[(1)] $\pi(1-\pi)^x$	$\frac{1}{\pi} - 1$	$\frac{1-\pi}{\pi^2}$	$\frac{\pi}{1-(1-\pi)e^t}$
Negative Binomial[(1)] $\binom{s+x-1}{x}\pi^s(1-\pi)^x$	$\frac{s}{\pi} - s$	$\frac{s(1-\pi)}{\pi^2}$	$\left[\frac{\pi}{1-(1-\pi)e^t}\right]^s$
Hypergeometric[(2)] $\frac{\binom{m}{x}\binom{n}{k-x}}{\binom{m+n}{k}}$	$N\pi$	$N\pi(1-\pi)\frac{N-k}{N-1}$	

Bernoulli Distribution

Interpolation

Perform one trial. Only two possible outcomes. Probability for success is p, for failure is q=1-p. x=1 means success, and x=0 means failure.

Definition

$$f_X(x) = \begin{cases} 1-p & \text{for } x=0 \\ p & \text{for } x=1 \end{cases}$$

Features

1. p is the parameter
2. E[X]=p
3. Var[X]=pq

Binomial Distribution

Interpolation

Perform n independent and identical Bernoulli trials with parameter p. X gives the total number of success in n trials.

Definition

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

Features

1. p, n are the parameters
2. $F_X(x) = \sum_{y=0}^{\lfloor x \rfloor} \binom{n}{y} p^y (1-p)^{n-y}$
3. E[X]=np
4. Var[X]=npq
5. $m_X(t) = (q + pe^t)^n, \quad m_X : \mathbb{R} \to \mathbb{R}$

Geometric Distribution

Interpolation

Perform a sequence of i.i.d. Bernoulli trials with parameter p, and stop until get one success. X gives the total number of trials needed to obtain the first success.

Definition

$$f_X(x) = (1-p)^{(x-1)}p$$

Features

1. p is the parameter
2. F(x)=$1 - q^{\lfloor x \rfloor}$
3. E[X]=$\frac{1}{p}$
4. Var[X]=$\frac{q}{p^2}$
5. $m_X(t) = \frac{pe^t}{1-qe^t}, \quad m_X : (-\infty, -\ln q) \to \mathbb{R}$

Pascal Distribution

Interpolation
Perform a sequence of i.i.d. Bernoulli trials with parameter p, and get the r^{th} success at the x^{th} trial.

Definition

$$f_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$$

Features

1. p, r are the parameters
2. E[X]=$\frac{r}{p}$
3. Var[X]=$\frac{rq}{p^2}$
4. $m_X(t) = \left(\frac{pe^t}{1-qe^t}\right)^r, \quad m_X : (-\infty, -\ln q) \to \mathbb{R}$

Connection to the Geometric Distribution

1. The Pascal distribution is a generalization of the Geometric distribution. Stop until get r success.
2. A random variable following the Pascal distribution with parameters r and p is the sum of r independent geometric random variables with parameter p.

Negative Binomial Distribution

Interpolation

Perform a sequence of i.i.d. Bernoulli trials with parameter p, and get the r^{th} success after x failures. The same as get the r^{th} success at the $(x+r)^{th}$ trial.

Definition

$$f_X(x) = \binom{x+r-1}{r-1} p^r (1-p)^x = \binom{-r}{x} (-1)^x p^r (1-p)^x$$

Features

1. p, r are the parameters
2. E[X]=$\frac{r(1-p)}{p}$. Can you imagine why?

The Negative Binomial

$$\binom{-r}{x} = \binom{r-1+x}{r-1} (-1)^x$$

Poisson Distribution

Interpolation

In a continuous interval [a, b], a certain event occurs for totally x times.

Assumptions

1. Independence: If the intervals T_1, $T_2 \subset$ [0, t] do not overlap (except perhaps at one point), then the numbers of arrivals in these intervals are independent of each other.
2. Constant rate of arrivals.

Definition

$$f_X(x) = \frac{k^x e^{-k}}{x!}$$

x = 0, 1, 2, 3, ...

Features

1. k is the parameter. $k = \lambda t$, where λ is the arrival rate and t is the length of the interval [a, b].
2. E[X]=k
3. Var[X]=k
4. $m_X(t) = e^{k(e^t-1)}$

Approximate the Binomial Distribution
A binomial distribution with large n and small p, can be approximated by a Poisson distribution with k=np, where you make sure the expected value is the same.
Because when $n \to \infty, n \cdot p = k$:

$$\binom{n}{x} p^x (1-p)^{n-x} \quad = \quad \frac{k^x}{x!} e^{-k}$$

Hypergeometric Distribution

Interpretation:
A total of N balls, r red and N − r black. Draw n balls out without putting back. Assume r > n and N − r > n. The random variable X describes the number of red balls in the n drawn balls.

Features:

1. N, r, n are the parameters
2. $f_X(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$
3. E[X]=$n\frac{r}{N}$
4. $\text{Var}\, X = n\frac{r}{N}\frac{N-r}{N}\frac{N-n}{N-1}$

Connection to the Bernoulli Distribution:
It is a sequence of identical but not independent Bernoulli trials. Each draw is a Bernoulli trial with $p = \frac{r}{N}$.

5 Continuous Random Variables

5.1 Definition

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

$$X : S \to \mathbb{R}$$

together with a probability **density** function

$$f_X : \mathbb{R} \to \mathbb{R}$$

where

(i) $f_X(x) \geq 0$ for all $x \in \mathbb{R}$

(ii) $\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1$

5.2 General Properties

1. f_X: probability density function(PDF). Notice in comparison to the discrete random variables, now $f_X \neq P[X = x]$. $P[X = x] = 0$.
2. $F_X(x) = \int_{-\infty}^{x} f_X(y)\,\mathrm{d}y$: cumulative distribution function(CDF). And $F'_X(x) = f_X(x)$ holds.
3. $\mathrm{E}[X] := \int_{-\infty}^{\infty} x \cdot f_X(x)\,\mathrm{d}x$: expectation.
4. $\mathrm{Var}[X] := \mathrm{E}\left[(X - \mathrm{E}[X])^2\right] = E[X^2] - E[X]^2$: Variance.
5. $m_X(t) := E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\,\mathrm{d}x$: moment generating function (MGF).

5.3 Locations

1. The median M_x: $P[X \leqslant M_x] = 0.5$
2. The mean $E[X]$: The average value.
3. The mode x_0: The location having the maximum f_X (if there is a unique maximum location).

5.4 Memoryless

Definition

$$P[X > x + s | X > x] = P[X > s]$$

Interpolation

Let's observe the definition above. What do you notice?
P[X>x+s | X>x]=P[X>s] does not rely on "x". We can interpret as:
P[X>x+s | X>x] does not "remember" that it is already with X>x.

5.5 Summary of continuous distributions

distribution pdf	mean	variance	mgf
Uniform $\begin{cases} \frac{1}{b-a} & \text{if } x \in [a,b] \text{ ,} \\ 0 & \text{otherwise} \end{cases}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$	$\begin{cases} \frac{e^{tb}-e^{ta}}{t(b-a)} & \text{if } t \neq 0 \text{ ,} \\ 1 & \text{if } t = 0. \end{cases}$
Standard normal $\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$	0	1	$e^{t^2/2}$
Normal $\frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$	μ	σ^2	$e^{\mu t + \sigma^2 t^2/2}$
Exponential $\lambda e^{-\lambda x}$	$1/\lambda$	$1/\lambda^2$	$\frac{\lambda}{\lambda - t}$
Gamma $\frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$	α/λ	α/λ^2	$\left[\frac{\lambda}{\lambda - t}\right]^\alpha$
Weibull $\frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha}$	$\beta\Gamma(1+\frac{1}{\alpha})$	$\beta^2 \left[\Gamma(1+\frac{2}{\alpha}) - [\Gamma(1+\frac{1}{\alpha})]^2\right]$	
Beta $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	

Exponential Distribution

Connection to the Poisson Distribution:
Start from the point of an arrival, the time for one successive arrival of a Poisson-distributed random variable to occur is exponentially distributed with parameter $\beta = \lambda$. (Recall: $k = \lambda t$)

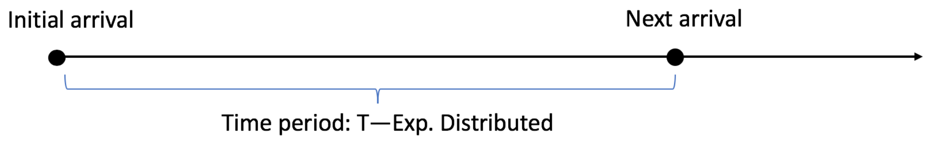


Figure 1: The exponential distribution

Definition:

$$f_\beta(x) = \begin{cases} \beta e^{-\beta x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Features:

1. β is the parameter.
2. $F_X(x) = 1 - e^{-\beta t}$
3. E[X]=$\frac{1}{\beta}$
4. Var[X]=$\frac{1}{\beta^2}$
5. $m_X(t) = (1 - \frac{t}{\beta})^{-1}$
6. Memoryless

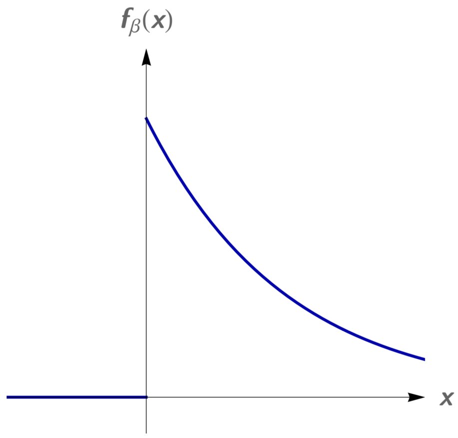


Figure 2: The exponential distribution

Gamma Distribution

Connection to the Poisson Distribution

Start from the point of an arrival, the time for r successive arrivals of a Poisson-distributed random variable to occur follows a gamma distribution with parameter $\alpha = r$, $\beta = \lambda$. (Recall: $k = \lambda t$)

Connection to the Exponential Distribution

It's a sum of i.i.d exponential Distributions.

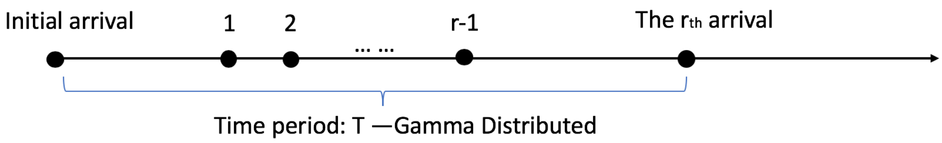


Figure 3: The gamma distribution

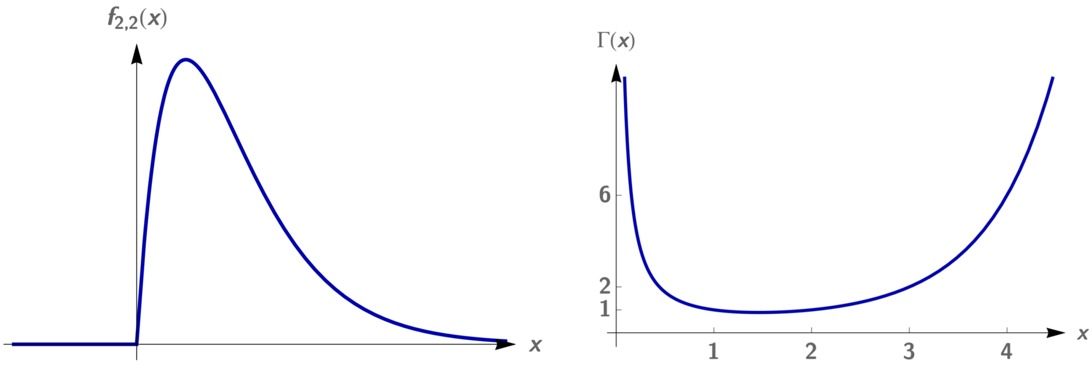
Definition

$$f_{\alpha,\beta}(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Features

1. α, β are the parameters.
2. $F_X(x) = 1 - e^{-\beta t}$
3. E[X]=$\frac{\alpha}{\beta}$
4. Var[X]=$\frac{\alpha}{\beta^2}$

5. $m_X(t) = (1 - \frac{t}{\beta})^{-\alpha}, \quad m_X : (-\infty, \beta) \to \mathbb{R}$
6. Not memoryless



(a) The gamma distribution (b) The gamma function

Figure 4: Gamma

Euler Gamma Function

1. $\Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z} dz$
2. $\Gamma(1) = 1$, $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ for $\alpha > 1$
3. $n! = \Gamma(n + 1)$, for $n \in \mathbb{R}$
4. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$
5. $\Gamma(\frac{2n+1}{2}) = \frac{(2n-1)(2n-3)\ldots 1}{2^n}\sqrt{\pi}$

Chi-squared Distribution

Connection to the Gamma Distribution

It's a special case of the gamma distribution, with $\alpha = \frac{\gamma}{2}$, $\beta = \frac{1}{2}$, where $\gamma \in \mathbb{N}$

Connection to the Normal Distribution

The sum of γ independent standard normal distribution follows a chi-squared distribution with γ degree of freedom.

Definition

$$f_\gamma(x) = \begin{cases} \frac{1}{\Gamma(\gamma/2)2^{\frac{\gamma}{2}}} x^{\gamma/2-1}e^{-x/2}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Features:

1. γ is the parameter, the degree of freedom
2. $F_X(x) = 1 - e^{-\beta t}$
3. E[χ^2_γ]=γ
4. Var[χ^2_γ]=2γ
5. $m_X(t) = (1 - \frac{t}{2})^{-\frac{\gamma}{2}}, \quad m_X : (-\infty, 2) \to \mathbb{R}$

Normal(Gau$\mathcal{B}$) Distribution

Definition

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-((x-\mu)/\sigma)^2/2}$$

Features

1. μ, δ are the parameters
2. E[X] = μ
3. Var[X] = δ^2
4. $m_X(t) = e^{\mu t + \sigma^2 t^2/2}$
5. $$P[-\sigma < X - \mu < \sigma] = 0.68$$ $$P[-2\sigma < X - \mu < 2\sigma] = 0.95$$ $$P[-3\sigma < X - \mu < 3\sigma] = 0.997$$

Standard Normal Distribution

Definition

Let X be a normally distributed random variable with mean μ and standard deviation δ. Then $Z = \frac{X-\mu}{\delta}$ follows a standard normal distribution with mean 0 and variance 1.

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

Features

1. E[Z] $= 0$
2. Var[Z] $= 1$
3. $m_Z(t) = e^{t^2/2}$

CDF

$$\Phi(z) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt = \frac{1}{2} erfc(-\frac{z}{\sqrt{2}})$$

Where we define:

$$\operatorname{erf}(z) := \frac{2}{\sqrt{\pi}} \int_{0}^{z} e^{-t^2} dt, \quad \operatorname{erfc}(z) := 1 - \operatorname{erf}(z)$$

Connection to the Chi-square Distribution

A Chi-square distributed variable with γ degree of freedom, is the sum of r square of standard normal distributed variables. Simply:

$$\chi_n^2 = \sum_{i=1}^{n} Z_i^2$$

Approximate the Binomial Distribution

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

Be careful with the half-unit correction.

Weibull Distribution

Definition

$$f(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}}, & x > 0 \\ 0, & \text{otherwise} \end{cases} \quad \alpha, \beta > 0$$

Features

1. $\mathrm{E}[X] = \alpha^{-1/\beta}\Gamma(1 + 1/\beta)$
2. $\operatorname{Var} X = \alpha^{-2/\beta}\Gamma(1 + 2/\beta) - \mu^2$

If f_A follows Weibull Distribution

1. $\varrho_A(t) = \alpha\beta t^{\beta-1}$
2. $R_A(t) = e^{-\alpha t^{\beta}}$

Connection to the Exponential Distribution
When $\beta = 1$, it is the exponential distribution.

Uniform Distribution

Definition

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{otherwise} \end{cases}$$

Features

1. $F(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & \text{otherwise} \end{cases}$
2. $\mathrm{E}[X] = \frac{a+b}{2}$
3. $\operatorname{Var} X = \frac{(b-a)^2}{12}$
4. $m_X(t) = \frac{e^{bt}-e^{at}}{t(b-a)}$

5.6 The Chebyshev Inequality

Theorem

Let c > 0 be any real number, and for k ∈ ℕ − 0, then for any random variables:

$$P[|X| \geq c] \leq \frac{E\left[|X|^k\right]}{c^k}$$

Application

In general, for any variables:

$$P[|X-\mu| \geq m\sigma] \leq \frac{1}{m^2}$$
$$P[-m\sigma < X-\mu < m\sigma] \geq 1-\frac{1}{m^2}$$

5.7 Transformation of Random Variables

Theorem

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

$$f_Y(y) = f_X\left(\varphi^{-1}(y)\right) \cdot \left|\tfrac{d\varphi^{-1}(y)}{dy}\right| \qquad \text{for } y \in \operatorname{ran}\varphi$$
$$f_Y(y) = 0 \qquad \text{for } y \notin \operatorname{ran}\varphi$$

It is important to know how to prove.

Proof

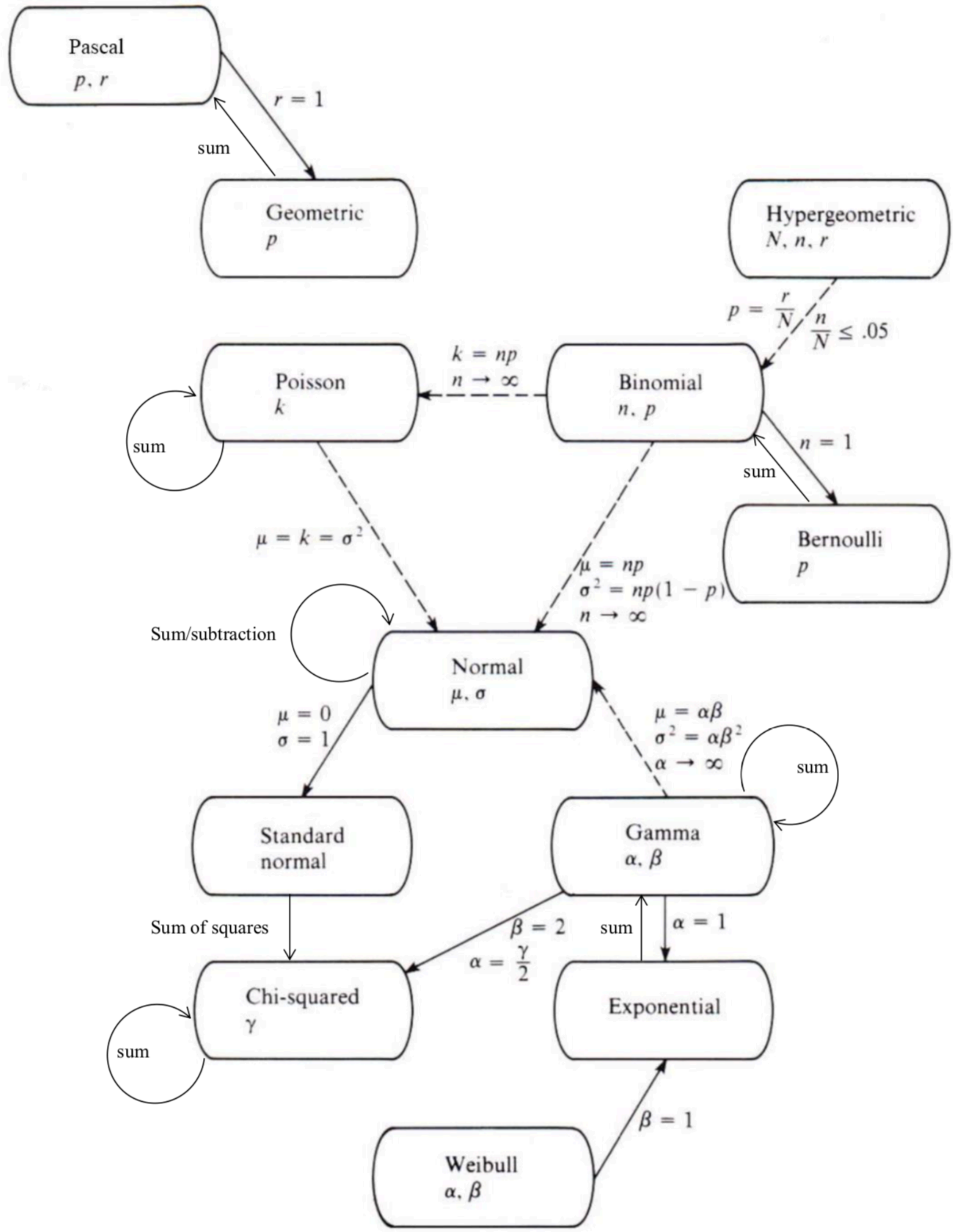
Step 1: Find CDF $F_Y(y)$.

$$\begin{aligned} F_Y(y) &= P[\varphi(X) \leq y] \\ &= P\left[\varphi^{-1}(\varphi(X)) \geq \varphi^{-1}(y)\right] \\ &= P\left[X \geq \varphi^{-1}(y)\right] \\ &= 1 - P\left[X \leq \varphi^{-1}(y)\right] \\ &= 1 - F_X\left(\varphi^{-1}(y)\right) \end{aligned}$$

Step 2: Find $f_Y(y) = F_Y'(y)$

$$\begin{aligned} f_Y(y) = F_Y'(y) &= -f_X\left(\varphi^{-1}(y)\right) \frac{d\varphi^{-1}(y)}{dy} \\ &= f_X\left(\varphi^{-1}(y)\right) \cdot \left|\frac{d\varphi^{-1}(y)}{dy}\right| \end{aligned}$$

6 Relationships of Distributions



7 Exercise

Exercise1 Integration

The distribution function of the speed (modulus of the velocity) V of a gas molecule is described by the Maxwell-Boltzmann law

$$f_V(v) = \begin{cases} \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{3/2} v^2 e^{-\frac{m}{kT}v^2/2} & v > 0 \\ 0 & v \leq 0 \end{cases}$$

where $m > 0$ is the mass of the molecule, $T > 0$ is its temperature and $k > 0$ is the Boltzmann constant.

i) Find the mean and variance of V.

ii) Find the mean of the kinetic energy $E = mV^2/2$.

iii) Find the probability density f_E of E.

(3 + 2 + 2 Marks)

Solution. i) The mean of V is given by

$$\begin{aligned} \mathrm{E}[V] &= \int_{\mathbb{R}} v f_V(v)\,dv \\ &= \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{1/2} \int_0^\infty \frac{m}{kT} v^3 e^{-\frac{m}{kT}v^2/2}\,dv \\ &= 2\left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{1/2} \int_0^\infty v e^{-\frac{m}{kT}v^2/2}\,dv \\ &= 2\left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{-1/2} [-e^{-\frac{m}{kT}v^2/2}]_0^\infty \\ &= 2\left(\frac{2kT}{m\pi}\right)^{1/2}. \end{aligned}$$

(1 Mark) Furthermore,

$$\begin{aligned} \mathrm{E}[V^2] &= \int_{\mathbb{R}} v^2 f_V(v)\,dv \\ &= \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{1/2} \int_0^\infty \frac{m}{kT} v^4 e^{-\frac{m}{kT}v^2/2}\,dv \\ &= 3\left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{1/2} \int_0^\infty v^2 e^{-\frac{m}{kT}v^2/2}\,dv \\ &= 3\left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{-1/2} \int_0^\infty e^{-\frac{m}{kT}v^2/2}\,dv. \end{aligned}$$

Setting $w = \sqrt{m/(kT)}v$, we have

$$\mathrm{E}[V^2] = 6\left(\frac{m}{kT}\right)^{-1} \underbrace{\frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-w^2/2}\,dw}_{=1/2} = \frac{3kT}{m}.$$

(1 Mark) It follows that

$$\operatorname{Var} V = \mathrm{E}[V^2] - \mathrm{E}[V]^2 = \frac{kT}{m}\left(3 - \frac{8}{\pi}\right)$$

(1 Mark)

ii) The expectation value of the kinetic energy is given by

$$\mathrm{E}[E] = \frac{m}{2}\mathrm{E}[V^2] = \frac{m}{2}\frac{3kT}{m} = \frac{3}{2}kT.$$

iii) Note that the function $\varphi\colon \mathbb{R} \to \mathbb{R}_+ \cup \{0\}$, $\varphi(v) = \frac{m}{2}v^2$ is not bijective, so we can't simply apply the theorem for transforming random variables from the lecture. Let $\varepsilon > 0$. Then

$$F_E(\varepsilon) = P[E \le \varepsilon] = P\left[\frac{m}{2}V^2 \le \varepsilon\right] = P\left[-\sqrt{2\varepsilon/m} \le V \le \sqrt{2\varepsilon/m}\right] = \int_{-\sqrt{2\varepsilon/m}}^{\sqrt{2\varepsilon/m}} f_V(v)\,dv$$
$$= \int_0^{\sqrt{2\varepsilon/m}} f_V(v)\,dx.$$

(1/2 Mark) It follows that

$$\begin{aligned} f_E(\varepsilon) = F_E'(\varepsilon) &= f_V(\sqrt{2\varepsilon/m}) \cdot \frac{1}{\sqrt{2m\varepsilon}} \\ &= \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{3/2} \frac{2\varepsilon}{m} e^{-\frac{\varepsilon}{kT}} \cdot \frac{1}{\sqrt{2m\varepsilon}} \\ &= \frac{2}{\sqrt{\pi}} (kT)^{-3/2} \sqrt{\varepsilon} e^{-\frac{\varepsilon}{kT}} \end{aligned}$$

for $\varepsilon > 0$. **(1 Mark)** For $\varepsilon \le 0$ we have

$$F_E(\varepsilon) = P[E \le \varepsilon] = P\left[\frac{m}{2}V^2 \le \varepsilon\right] \le P\left[\frac{m}{2}V^2 \le 0\right] = 0,$$

so $f_E(\varepsilon) = 0$ for $\varepsilon \le 0$. **(1/2 Mark)**

Sample2 Relationships

Exercise 2.1
Let X be a discrete random variable following a Bernoulli distribution with parameter $p = 1/2$ and let $X_1, \ldots, X_{10}$ be a random sample of size 10. Calculate the probability that the sample mean is greater than 3/4, i.e., find

$$P[\overline{X} > 3/4].$$

Solution. We note that $X_1 + \cdots + X_{10}$ follows a binomial distribution with $n = 10$ and $p = 1/2$.

Exercise 2.2
Let X be a discrete random variable following a geometric distribution with parameter $p = 1/2$ and let $X_1, \ldots, X_{10}$ be a random sample of size 10. Calculate the probability that the sample mean is no more than than 1.5, i.e., find

$$P[\overline{X} \le 1.5].$$

Solution. We note that $X_1 + \cdots + X_{10}$ follows a Pascal distribution with $r = 10$ and $p = 1/2$.

Exercise 2.3
Let X be a discrete random variable following a Poisson distribution with parameter $k = 2$ and let $X_1, \ldots, X_{10}$ be a random sample of size 10. Calculate the probability that the sample mean is no more than than 1.5, i.e., find

$$P[\overline{X} \le 1.5].$$

Solution. We note that $X_1 + \cdots + X_{10}$ follows a Poisson distribution with $k = 20$.

Sample4 Conditional Probability

Exercise 4.2

A company produces toy plastic coins for use in board games. They will be tossed and should return either "heads" or "tails" withe qual probability $p_0 = 0.5$. Most of the coins are fine, but due to a molding process fault, 5% of the coins are defective and have a $p = 0.7$ chance of returning "heads".

A coin is tested by tossing it 100 times and recording the number of heads. It will be deemed defective and discarded if "heads" occurs at least 70 times.

Given that a coin is discarded, what is the probability that it was defective?

Solution. We know that $P[\text{defective}] = 0.05$. Furthermore,

$$P[\text{discarded} \mid \text{not defective}] = \frac{1}{2^n} \sum_{x=70}^{100} \binom{100}{x} = 0.000039$$

and

$$P[\text{discarded} \mid \text{defective}] = \sum_{x=70}^{100} \binom{100}{i} 0.7^x 0.3^{100-x} = 0.549$$

Then

$$\begin{aligned} P[\text{defective} \mid \text{discarded}] &= \frac{P[\text{discarded} \mid \text{defective}] \cdot P[\text{defective}]]}{P[\text{discarded} \mid \text{defective}] \cdot P[\text{defective}]] + P[\text{discarded} \mid \text{not defective}] \cdot P[\text{not defective}]]} \\ &= \frac{0.549 \cdot 0.05}{0.549 \cdot 0.05 + 0.000039 \cdot 0.95} \\ &= 0.999 \end{aligned}$$

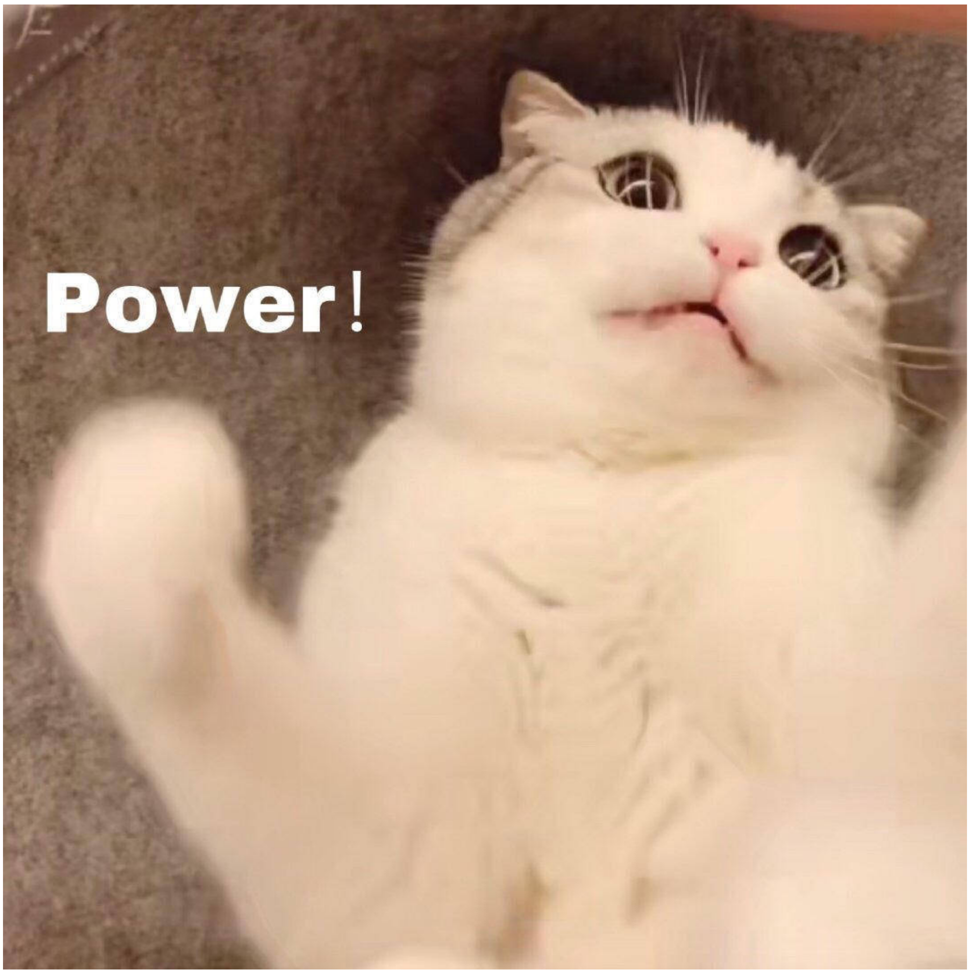
i) *2 Marks for writing down the correct probabilities based on the exercise description.*

ii) *2 Marks for the calculation using conditional probabiliy/Bayes's theorem.*

iii) *1 Mark for the correct correct result, if it is supported by calculation above.*

8 Tips for Mid

1. Know how to integration.
2. Understand why and when you can use the methods and theorems.
3. Understand how to calculate and use E[X], Var[X], MGF, CDF, ... for arbitrary distributions, instead of memorizing or searching them for common distributions.
4. Understand the meaning and relationships between different distributions, instead of memorizing these equations for a single distribution.
5. Understand the parts related to the conditional probability.
6. Prove any statement not proved in class or in homeworks. Although probably you do not need such statements.



Wish you good luck in all of your midterm exams!