Contents

1

Definitions and Theory

1.1 Foundations

Definition 1. *Two random vectors* $(X_1, X_2, ..., X_n)$ *and* $(Y_1, Y_2, ..., Y_n)$ *are said to be independent if* $F(x_1, x_2, ..., x_m, y_1, y_2, ..., y_n) = F_1(x_1, x_2, ..., x_m)F_2(x_1, x_2, ..., x_n)$ *for all* $(x_1, x_2, ..., x_m, y_1, y_2, ..., y_n) \in \mathbb{R}^{m+n}$ *where* F, F_1, F_2 *are the joint CDF's of* $(X_1, X_2, ..., X_m, Y_1, Y_2, ..., Y_n)$, $(X_1, X_2, ..., X_m)$ *and* $(Y_1, Y_2, ..., Y_n)$ *respectively.*

Theorem 1. *Let* $X = (X_1, X_2, , X_m)$ *and* $Y = (Y_1, Y_2, , Y_n)$ *be independent random vectors. Then the component* X_j *of* $X(j = 1, 2, , m)$ *and the component* Y_k *of* $Y(k = 1, 2, , n)$ *are independent random variables. If* h *and* g *are Borel-measurable functions,* $h(X_1, X_2, , X_m)$ *and* $g(Y_1, Y_2, , Y_n)$ *are independent.*

Theorem 2. *Let* X *be a random variable with pdf* $f(x)$*. Then the pdf of* $aX + b$ *where* $a \neq 0, b \in \mathbb{R}$ *is given by* $\frac{1}{a} f\left(\frac{x-b}{a}\right)$

Theorem 3. *If* X *and* Y *are independent continuous random variables with pdfs* $f_X(x)$ *and* $f_Y(y)$*, then the pdf of* $Z = X+Y$ *is* $f_Z\left(z\right) = \int_{-\infty}^{\infty} f_X\left(w\right) f_Y\left(z-w\right) \mathrm{d}w$

Definition 2. *The* ***expectation*** *of a random variable* X *having values* $x_1, x_2, ..., x_n$ *is defined as* $E\left(X\right) = x_1 P\left(X = x_1\right) + x_2 P\left(x = x_2\right) + ... + x_n P\left(X = x_n\right) = \sum_{j=1}^{n} x_j P\left(X = x_j\right) = x_1 f\left(x_1\right) + x_2 f\left(x_2\right) + ... + x_n f\left(x_n\right) = \sum_{j=1} \to n x_j f\left(x_j\right)$ *where* f *is the distribution of* X*.*

Theorem 4. *The expectation has the following properties:*

1. $E(cX) = cE(X)$ *where* c *is any constant*

2. *If* X *and* Y *are any random variables then* $E(X+Y) = E(X) + E(Y)$

3. *If* X *and* Y *are independent random variables, then* $E(XY) = E(X)E(Y)$

Definition 3. *The **variance** is defined as* $Var(X) = \sigma^2 = E[\ (X-\mu)^2\ ]$. *The **standard deviation** is defined as* $\sigma = \sqrt{Var(X)}$.

Theorem 5. *The variance has the following properties:*

1. $\sigma^2 = E\left[(X - m_u)^2\right] = E\left(X^2\right) - \mu^2 = E\left(X^2\right) - \left[E\left(X\right)\right]^2$ *where* $\mu = E\left(X\right)$

2. *If* c *is any constant,* $Var(cX) = c^2\, Var(X)$

3. *The quantity* $E[(X-a)^2]$ *is a minimum where* $a = \mu = E(X)$

4. *If* X *and* Y *are independent random variables*

Definition 4 (Moments)**.** *The* r*-th **moment** of a random variable* X *about the mean* μ *is defined as* $\mu_r = E[\ (X-\mu)^r\ ]$ *where* $r = 0, 1, 2, ...$
The r*-th moment of* X *about the origin is defined as* $\mu_r' = E(X^r)$.

It follows that $\mu_0 = 0$, $\mu_1 = 1$, $\mu_2 = \sigma^2$

Definition 5 (Moment Generating Function)**.** *Let* X *be a random variable defined on* $(\Omega, \mathcal{F}, P)$. *The function* $M(t) = E\left[e^{tX}\right]$ *is called the **moment generating function** (MGF) of the random variable* X *if the expectation on the right side exists in some neighbouhood of the origin. If the expectation on the right side does not exist in any neighbourhood of the origin, then we say the MGF does not exist.*

The r-th derivative of the moment generating funtion is the r-th moment about the origin μ_r'.

Theorem 6. *If the MGF* $M(s)$ *of a random variable* X *exists, then the MGF* $M(s)$ *has derivatives of all orders at* $s = 0$ *and*

$M^{(k)}(s)|_{s=0} = EX^k$ *for positive integer* k

Theorem 7. *The moment generating function has the following properties:*

1. *For any constants* a *and* b*, the mgf of the random variable* $aX + b$ *is given by* $M_{aX+b} = e^{bt} M_X(at)$

2. *If* X *and* Y *are independent random variables having moment generating functions* $M_X(t)$ *and* $M_Y(T)$ *respectively, then* $M_{X+Y}(t) = M_X(t)M_Y(t)$

3. ***Uniqueness Theorem*** *Suppose that X and Y are random variables having moment generating functions $M_X(t)$ and $M_Y(t)$ respectively. Then X and Y have the same probability distribution if and only if $M_X(t) = M_Y(t)$ identically.*

Definition 6. *Let $X_1, X_2, ..., X_n$ be a jointly distributed or $(X_1, X_2, ..., X_n)$ be a random vector. If $E[\exp(\sum_{j=1}^{n} t_j X_j)]$ exists for $|t_j| \leq h_j$, $j = 1, 2, ..., n$, we write $M(t_1, t_2, ..., t_n) = E[\exp(t_1 X_1 + t_2 X_2 + ... + t_n X_n)]$ and call it the MGF of $X_1, X_2, ..., X_n$ or simply, the MGF of the random vector $(X_1, X_2, ..., X_n)$*

Theorem 8. *The joint MGF $M(t_1, t_2)$ uniquely determines the joint distribution of (X, Y). Conversely, if the joint MGF exists it is unique.*

Theorem 9. *The joint MGF $M(t_1, t_2)$ completely determines the marginal distributions of X and Y.*
$M(t_1, 0) = E[\exp(t_1 X)] = M_X(t_1)$ and $M(0, t_2) = E[\exp(t_2 X)] = M_Y(t_2)$

Theorem 10. *X and Y are independent random variables if and only if $M(t_1, t_2) = M(t_1, 0)M(0, t_2)$ for all $t_1 \in [-h_1, h_1], t_2 \in [-h_2, h_2]$*

1.2 Distributions

Definition 7 (Cauchy Distribution)**.** *A variable the pdf $f(x) = \dfrac{1}{\sigma\pi(1 + \left(\frac{x-\mu}{\sigma}\right)^2)}$, $x \in \mathbb{R}$ where $\sigma > 0, \mu \in \mathbb{R}$ is called a **Cauchy random variable** with parameter μ and σ^2. We write $X \sim \mathcal{C}(\mu, \sigma^2)$ for Cauchy random X with pdf.*

Definition 8 (Gamma Function)**.** *$\Gamma(\alpha)$ denotes the gamma function which is defined as $\Gamma(\alpha) = \displaystyle\int_0^\infty t^{\alpha-1} e^{-t}\, \mathrm{d}t$, where $\alpha > 0$.*

Theorem 11 (Properties of the Gamma function)**.**

1. *The improper integral converges for all $\alpha > 0$.*
2. *$\Gamma(\alpha) > 0$ for all $\alpha > 0$, $\Gamma(\alpha) \to \infty$ as $\alpha \to 0$.*
3. *$\Gamma(1) = 1$*
4. *$\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$*
5. *For $n \in \mathbb{N}$, $\Gamma(n) = (n-1)!$*
6. *$\Gamma(\frac{1}{2}) = \sqrt{\pi}$*

Definition 9 (Gamma Distribution)**.** *A random variable with the pdf* $f(x) = \dfrac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$ *where* $0 < x < \infty, \alpha > 0, \beta > 0$ *is called a* ***gamma random variable*** $G(\alpha, \beta)$ *with parameters* α *and* β*.*

$G(1, \beta) = \exp(\beta)$ with pdf $f(x) = \frac{1}{\beta} e^{-x/\beta}$, $x > 0$.

Definition 10 (Chi-Square Distribution)**.** *The* ***chi-square distribution*** *is* $f(x) = \dfrac{1}{\Gamma(\frac{p}{2}) 2^{\frac{p}{2}}} x^{\frac{p}{2}-1} e^{-x/2}$ *where* $0 < x < \infty$*.*

χ^2_p *denotes a chi-square random variable with* p *degrees of freedom.*

The chi-square distribution is a special case of the gamma distribution with $\alpha = \frac{p}{2}$ and $\beta = 2$.
Let $X \sim N(0, 1)$. Then $X^2 \sim \chi^2_1$

1.3 Sampling Theory

We can either sample *with replacement* or *without replacement*. A finite population sampled with replacement can be considered infinite. Sampling from a very large finite population can similarly be considered as sampling from an infinite population.
To properly choose the sample, we can make sure that every member of the population has an equal chance of being in the sample. Normally, since the sample size is much smaller than the population size, sampling without replacement will give practically the same results as sampling with replacement.
For a sample of size n from a population which we assume has distribution $f(x)$, we can choose members of the population at random, each selection corresponding to a random variable $X_1, X_2, ..., X_n$ with corresponding values $x_1, x_2, ..., x_n$. In case we are assuming sampling without replacement, $X_1, X_2, ..., X_n$ will be independent and identically distributed random variables with probability distribution $f(x)$.

Definition 11 (Random Sample)**.** *Let* X *be a random variable with a distribution* f*, and let* $X_1, X_2, ..., X_n$ *be iid random variables with the common distribution* f*. Then the collection* $X_1, X_2, ..., X_n$ *is called a* ***random sample*** *of size* n *from the population* f*.*

Since $X_1, X_2, ..., X_n$ are iid, the joint distribution of the random sample is $f(x_1, x_2, ..., x_n) = f(x_1)f(x_2)...f(x_n)$.
Any quantity obtained from a sample for the purpose of estimating a population parameter is called a sample statistic, or briefly statistic. Mathematically, a sample statistic for a sample of size n can be defined as a function of the random variables $X_1, X_2, ..., X_n$ as $T(X_1, X_2, ..., X_n)$. This itself is a random variable whose values can be represented as $T(x_1, x_2, ..., x_n)$.

Definition 12 (Statistic)**.** *Let $X_1, X_2, ..., X_n$ be a random sample of size n from the population whose distribution is $f(x|\theta)$ (the distribution f with unknown parameter θ). Let $T(x_1, x_2, ..., x_n)$ be a real-valued or vector-valued function whose domain includes the range of $(X_1, X_2, ..., X_n)$. Then the random variable or random vector $Y = T(X_1, ..., X_n)$ is called a* ***statistic*** *provided that T is not a function of any unknown parameter θ.*

For example consider $X \approx N(\mu, \sigma^2)$ where μ is known but σ is unknown. Then $\frac{\sum_{i=1}^{n} X_i}{\sigma^2}$ is not a statistic but $\frac{\sum_{i=1}^{n} X_i}{\mu^2}$ is a statistic.
Two common statistics are the sample mean and sample variance.

Definition 13 (Sample Mean, Sample Variance, Sample Standard Deviation)**.** *The* ***sample mean*** *is the arithmetic average of the values in the random sample. It is denoted by $\bar{X} = \dfrac{X_1 + X_2 + ... + X_n}{n}$.*

The ***sample variance*** *is the statistic defined by $S^2 = \dfrac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2}{n-1}$*
The ***sample standard deviation*** *is the statistic defined by $S = \sqrt{S^2}$.*

Definition 14 (Unbiased Estimator)**.** *Let $X_1, X_2, ..., X_n$ be a random sample from a population $f(x|\theta)$. We say that a statistic $T(X_1, X_2, ..., X_n)$ is an* ***unbiased estimator*** *of the parameter θ if $E_T = \theta$ for all possible values of θ.*

Theorem 12. *Let $X_1, X_2, ..., X_n$ be a random sample from a population with mean μ and variance $\sigma^2 < \infty$. Then:*

1. $E\left(\bar{X}\right) = \mu$
2. $Var\left(\bar{X}\right) = \frac{\sigma^2}{n}$
3. $E\left(S^2\right) = \sigma^2$

From the above theorem we see that the sample mean $\bar{X}$ is an unbiased estimator of the population mean μ and the sample variance $\S^2$ is an unbiased estimator of the population variance σ^2. (The reason we included $1/n-1$ in the definition of the sample variance was to make it an unbiased estimator)

Definition 15. *Let $X_1, X_2, ..., X_n$ be a random sample of size n from a population $f(x|\theta)$. The probability distribution of a statistic $T(X_1, X_2, ..., X_n)$ is called the sampling distribution of T.*

Theorem 13 (Distribution of the sample mean)**.** *Let $X_1, X_2, ..., X_n$ be a random sample from a population with MGF $M_X(t)$. Then the MGF of the sample mean is $M_{\bar{X}}(t) = (M_X(t/n))^n$.*

Theorem 14. *Let $X_1, X_2, ..., X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution, and let X denote the sample mean.*
Then X and the random vector $(X_1 - X, X_2 - X, ..., X_n - X)$ are independent.

Theorem 15. *Let $X_1, X_2, ..., X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution, and let X denote the sample mean and S^2 denote the sample variance. Then X and S^2 are independent random variables.*

The converse of this theorem is also true: if the sample mean and sample variance of a random sample are independent random variables then population distribution is normal.

2

Exercises

1.

(i) Let X be a standard normal variable $N(0,1)$. Find the moment generating function of X.

(ii) Let Y be a standard normal variable $N(\mu, \sigma^2)$. Find the moment generating function of Y.

ans.
$M_X(t) = e^{\frac{1}{2}t^2}$
$M_Y(t) = e^{\mu t}e^{\frac{\sigma^2 t^2}{2}}$

Source. Class, Lec 04

2. Let X have the PMF $P(X = k) = \frac{6}{\pi^2 k^2}$ where $k = 1, 2, ...$ Show that the MGF of X does not exist.

Source. Class, Lec 04

3. Let $X \approx \exp(\lambda)$ with pdf $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$. Then show that $E[e^{tX}]$ exists for all $t < \lambda$.

Source. Class, Lec 04

4. Let $Z_1, Z_2, ..., Z_n$ be a random sample from $\mathcal{C}(0,1)$ distribution. Derive the distribution of sample mean $\bar{Z}$.

ans.
$Cauchy(0,1)$

Source. Class, Lec 05

5. Show that the MGF of the gamma distribution $G(\alpha, \beta)$ is $M_G = \dfrac{1}{(1-\beta t)^\alpha}$ for $t < \dfrac{1}{\beta}$. Hence compute the mean and variance of the gamma distribution.

Source. Class, Lec 06

6. Let $X_1, X_2, ..., X_n$ be independent random variables such that $X_j \sim G(\alpha_j, \beta)$, $j = 1, 2, ..., n$. Then show that $X_1 + X_2 + ... + X_n \sim G(\sum_{j=1}^{n} \alpha_j, \beta)$.

Source. Class, Lec 06

7. Let $X \sim N(0, 1)$. Then show that $X^2 \sim \chi_1^2$

Source. Class, Lec 06

8. Let $X \sim U(0, 1)$. Then show that $-2 \ln X \sim \chi_2^2$

Source. Class, Lec 06

9. Let $X_1, X_2, ..., X_n$ be a random sample of n identical circuit boards whose times until failure are thought to follow an exponential(β) population. Find the joint distribution of the sample. What is the probability that all the boards last more than 2 years?

ans.
$f(x_1, x_2, ..., x_n) = f(x_1)f(x_2)...f(x_n) = \Pi_{i=1}^{n} \frac{1}{\beta} \exp(\frac{-x_i}{\beta})$
$P(X_1 > 2, X_2 > 2, ..., X_n > 2) = \exp(\frac{-2n}{\beta})$

Source. Class, Lec 02

10. If sample $X_1, X_2, ..., X_n$ are drawn from a finite population without replacement, then show that the random variables $X_1, X_2, ..., X_n$ are not mutually independent but that they are identically distributed.

Source. Class, Lec 02

11. Let $X \approx$ Bernoulli (p) where p is possibly unknown. Suppose that five independent observations on X are $0, 1, 1, 1, 0$ Then find the sample mean, sample variance and sample standard deviation.

ans.
Mean $\bar{x} = 0.6$
Variance $s^2 = 0.3$
Standard Deviation $s = 0.55$

Source. Class, Lec 03

12. Let $X_1, X_2, ..., X_n$ be a random sample from a population with mean μ and variance $\sigma^2 < \infty$. Then show that:

1. $E\left(\bar{X}\right) = \mu$
2. $\operatorname{Var}\left(\bar{X}\right) = \frac{\sigma^2}{n}$
3. $E\left(S^2\right) = \sigma^2$

Source. Class, Lec 03

13. Color blindness appears in 1% of the people in a certain population. How large must a sample be if the probability of its containing a color-blind person is to be 0.95 or more? (Assume that the population is large enough to be considered infinite, so that sampling can be considered to be with replacement.)

ans.
Sample size greater than 299

Source. Class, Lec 03

14. Suppose that five observations on normal population are $-0.864, 0.561, 2.355, 0.582, -0.774$. Compute sample variance.

ans.
1.648

Source. Class, Lec 03

15. For a sample of size 5 that results in 7,9,1,6,2, find the sample mean and variance.

ans.
Sample Mean $\bar{x} = 6$ *Sample Variance* $S^2 = 2.5$

Source. Schaum's Example 5.5

16. Let $X_1, X_2, ..., X_n$ be a random sample from a population with MGF $M_X(t)$. Then show that the MGF of the sample mean is $M_{\bar{X}}(t) = (M_X(t/n))^n$.

Source. Class, Lec 04

17. Let $X_1, X_2, ..., X_n$ be a random sample from a $N(\mu, \sigma^2)$ population. Find the distribution of the sample mean.

ans.
$\exp(\mu t + \frac{(\sigma^2/n)t^2}{2})$, *i.e.* $\bar{X}$ *has a* $N(\mu, \frac{\sigma^2}{n})$ *distribution.*

Source. Class, Lec 04

18. Let $X \sim (C)(0, 1)$. Show that:

1. EX does not exist
2. the mean does not exist
3. the MGF does not exist

Source. Class, Lec 04

19. Let $X_1, ..., X_n$ be iid random variables with continuous cdf F_X and suppose $EX_i = \mu$. Define the random variables $Y_1, Y_2, ..., Y_n$ by $Y_i = \begin{cases} 1 & \text{if } X_i > \mu \\ 0 & \text{if } X_i \le \mu \end{cases}$. Find the distribution of $\sum_{i=1}^{n} Y_i$

ans.
Binomial$(n, 1 - F_X(\mu))$

Source. Class, Lec 04

20. A fair die is rolled. Let X be the face value that turns up, and X_1, X_2 be two independent observations on X. Compute the PMF of $\bar{X}$.

Source. Class, Lec 04

21. Let $X_1, X_2, ..., X_n$ be independent observations on Poisson population with parameter λ. Find the distribution of the sample mean.

ans.
$$P(\bar{X} = t) = \frac{e^{-n\lambda}(n\lambda)^{tn}}{(tn)!} \text{ where } t = 0, \tfrac{1}{n}, \tfrac{2}{n}, ...$$

Source. Class, Lec 04

22. Let $X_1, X_2, ..., X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution, and let X denote the sample mean.
Then show that X and the random vector $(X_1 - X, X_2 - X, ..., X_n - X)$ are independent.

Source. Class, Lec 06

23. Let $X_1, X_2, ..., X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution, and let X denote the sample mean and S^2 denote the sample variance. Then show that X and S^2 are independent random variables.

Source. Class, Lec 06