

Math 324 Final Review

①

Part I

- Joint distributions

- Discrete

$$p(x, y) = P(X=x \text{ and } Y=y)$$

- Continuous

$$f(x, y)$$

- Probabilities

$$\text{Discrete: } P[(X, Y) \in A] = \sum_{(x, y) \in A} p(x, y)$$

$$\text{Continuous: } P[(X, Y) \in A] = \iint_A f(x, y) dy dx$$

- Marginal probabilities

$$P_X(x) = \sum_y p(x, y)$$

$$P_Y(y) = \sum_x p(x, y)$$

$$f_X(x) = \int f(x, y) dy$$

$$f_Y(y) = \int f(x, y) dx$$

- independence

$$p(x, y) = p(x) p(y)$$

$$X, Y \text{ independent} \iff f(x, y) = f(x) f(y)$$

- Conditional Distributions

$$f_{y|x}(y|x) = \frac{f(x, y)}{f_x(x)}$$

- Expected value

$$E[h(X, Y)] = \begin{cases} \sum_x \sum_y h(x, y) p(x, y) \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) dy dx \end{cases}$$

- Covariance

$$\text{cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)]$$

$$= \begin{cases} \sum_{x, y} (x - \mu_x)(y - \mu_y) p(x, y) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) f(x, y) dy dx \end{cases}$$

$$\text{cov}(X, Y) = E[XY] - E[X]E[Y]$$

- Correlation

$$\rho_{x,y} = \frac{\text{Cov}(X,Y)}{\sigma_x \sigma_y}$$

X, Y independent $\Rightarrow \rho = 0$ (but reverse not true)

$$-1 \leq \rho \leq 1$$

- Sampling Distributions

- statistic

- quantity estimated from data
- sampling variability

- Random samples

- i.i.d

- Finding sampling distribution

- Theory
- Simulation

④
- distribution of \bar{X} and T_n

X_1, \dots, X_n iid $E[X_i] = \mu$
 $Var(X_i) = \sigma^2$

$$- E[\bar{X}] = \mu$$

$$- Var[\bar{X}] = \frac{\sigma^2}{n}$$

$$- E[T_n] = n\mu$$

$$- Var(T_n) = n\sigma^2$$

- If X_1, \dots, X_n iid $N(\mu, \sigma^2)$

\bar{X} exactly ~~is~~ $N(\mu, \frac{\sigma^2}{n})$

T_n exactly $N(n\mu, n\sigma^2)$

- CLT if $n \rightarrow \infty$ then

\bar{X} is $N(\mu, \sigma^2)$ approximately

T_n is $N(n\mu, n\sigma^2)$ approximately

(rule of thumb $n > 30$)

- linear combinations

$$a_1x_1 + \dots + a_nx_n = \sum a_i x_i$$

$$E[\sum a_i x_i] = \sum a_i E[x_i]$$

$$Var(\sum a_i x_i) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j Cov(x_i, x_j)$$

Note this implies that if x, y independent

$$Var(x_1 - x_2) = Var(x_1) + Var(x_2)$$

- if x_1, \dots, x_n are all Normal then

$\sum a_i x_i$ is also normal

- Point estimation - population characteristic
- parameter - single number
- point estimate - single number
- point estimator - method or formula

eg	μ	\bar{X}
	σ	S
	p	\hat{p}

- Bias

$$\text{Bias}(\hat{\theta}) = E[\hat{\theta}] - \theta$$

$$\text{Unbiased} \Rightarrow E[\hat{\theta}] = \theta$$

$$E[\bar{X}] = \mu$$

$$E[S^2] = \sigma^2$$

$$E[\hat{\rho}] = \rho$$

} All unbiased

- unbiased estimators are better

- Minimum Variance Unbiased Estimation

- choose estimator with smallest variance

- If X_1, \dots, X_n from $N(\mu, \sigma^2)$ then

\bar{X} is MVUE for μ .

- Standard error - (estimate of the standard deviation of an estimator)

$$\text{eg } SE[\bar{X}] = \frac{s}{\sqrt{n}}$$

$$SE(\hat{\rho}) = \sqrt{\frac{\hat{\rho}(1-\hat{\rho})}{n}}$$

- MME (Method of Moments estimation) (7)

equating

$$E[X^k] = \frac{\sum X^k}{n} \quad \text{for } k=1, 2, \dots$$

Solve system of equations

- MLE (Maximum likelihood estimation)

$$\nearrow L(\hat{\theta}) = f(x_1, x_2, \dots, x_n; \hat{\theta}) \quad \hat{\theta} = (\theta_1, \dots, \theta_p)$$

likelihood

log likelihood

$$l(\theta) = \log L(\hat{\theta})$$

usually differentiate w.r.t to each of $\theta_1, \dots, \theta_p$

Solve system of equations i.e.

$$\frac{\partial l}{\partial \theta_1} = 0$$

$$\frac{\partial l}{\partial \theta_p} = 0$$

- IF $\hat{\theta}_1, \dots, \hat{\theta}_p$ is MLE for $\theta_1, \dots, \theta_p$ (9)

Then $h(\hat{\theta}_1, \dots, \hat{\theta}_p)$ is MLE of $h(\theta_1, \dots, \theta_p)$

- when n becomes very large

$\hat{\theta}$ becomes unbiased

$\hat{\theta}$ has minimum variance