CS 3341 – Probability and Statistics

Cheat Sheet

Ian Macalinao

Chapter 3 – Discrete Random Variables And Their Distributions

Bernoulli distribution

- p probability of success
- $\mathbf{E}(X) = p$
- Var(X) = pq

Binomial distribution

- n number of trials
- p probability of success
- q probability of failure
- $P(x) = \binom{p}{x} p^x q^{n-x}$
- $\mathbf{E}(X) = np$
- Var(X) = npq

Geometric distribution

- \bullet p probability of success
- $P(x) = (1-p)^{x-1}p, x = 1, 2, ...$
- $\mathbf{E}(X) = \frac{1}{p}$
- $\operatorname{Var}(X) = \frac{1-p}{p^2}$

Negative Binomial distribution

- k number of successes
- p probability of success

- $P(x) = {x-1 \choose k-1} (1-p)^{x-k} p^k, x = k, k+1, ...$ $\mathbf{E}(X) = \frac{k}{p}$ $Var(X) = \frac{k(1-p)}{p^2}$

Poisson distribution

- λ frequency, average number of events
- $P(x) = e^{-\lambda} \frac{\lambda^x}{x!}, x = 0, 1, 2, ...$ $\mathbf{E}(X) = \lambda$
- $Var(X) = \lambda$

Chapter 4 – Continuous Distributions

• $\mathbf{E}(X) = \int_0^\infty x f(x) dx$

Uniform distribution

- $f(x) = \frac{1}{b}$ for $0 \le x \le b$. $\mathbf{E}(X) = \frac{a+b}{2}$ $Var(X) = \frac{(b-a)^2}{12}$

Exponential distribution

- λ frequency parameter, the number of events per time unit
- $\mathbf{E}(x) = \frac{1}{\lambda}$ $Var(X) = \frac{1}{\lambda^2}$

Gamma distribution

- α shape parameter
- λ frequency parameter
- $\mathbf{E}(X) = \frac{\alpha}{\lambda}$ $\mathrm{Var}(X) = \frac{\alpha}{\lambda^2}$

Gamma-Poisson formula

- $\mathbf{P}\{T > t\} = \mathbf{P}\{X < \alpha\}$ $\mathbf{P}\{T \le t\} = \mathbf{P}\{X \ge \alpha\}$

Normal distribution

- μ expectation, location parameter
- σ standard deviation, scale parameter
- $\mathbf{E}(X) = \mu$
- $Var(X) = \sigma^2$
- $Z = \frac{x \mu}{\sigma}$ $Z_n = \frac{S_n n\mu}{\sigma\sqrt{n}}$ Central limit theorem

Normal can be approximated to Binomial with Binomal $(n, p) \approx \text{Normal}(\mu =$ $np, \sigma = \sqrt{np(1-p)}$

Chapter 6 – Stochastic processes

Binomial counting process

- $\lambda = \frac{p}{\Delta}$ $n = \frac{t}{\Delta}$ X(n) = Binomial(n, p) = np• $Y = \text{Geometric}(p) = \frac{1}{p}$
- $T = Y\Delta$
- $\mathbf{E}(T) = \mathbf{E}(Y)\Delta = \frac{\Delta}{p} = \frac{1}{\lambda}$ $Var(T) = Var(Y)\Delta^2 = (1-p)(\frac{\Delta}{p})^2 = \frac{1-p}{\lambda^2}$

Poisson process

- $\mathbf{E}X(t) = \lambda t$
- $X(t) = Poisson(\lambda t)$
- $T = \text{Exponential}(\lambda)$

- T_k = Gamma(k, λ)
 P{T_k ≤ t} = P{X(t) ≥ k}
 P{T_k > t} = P{X(t) < k}

Chapter 7 – Queuing Systems

Parameters of a queuing system

- λ_A = arrival rate, i.e. the expected arrivals per 1 unit of time.
- λ_S = service rate, the average number of jobs processed by a continuously working server during one unit of time.
- $\mu_A = 1/\lambda_A$ mean interarrival time

- $\mu_S = 1/\lambda_S$ mean service time
- $r = \lambda_A/\lambda_S = \mu_S/\mu_A$ = utilization, or arrival-to-service ratio, or the expected number of jobs receiving service at any given time

Random values of a queuing system

- $X_s(t)$ = number of jobs receiving service at time t
- $X_w(t)$ = number of jobs waiting in a queue at time t
- $X(t) = X_s(t) + X_w(t)$ = the total number of jobs in the system at time t. This is a **queueing process** because it may increase or decrease, whereas a **counting process** may only increase.
- S_k = service time of the k-th job
- W_k = waiting time of the k-th job
- $R_k = S_k + W_k$ = response time, the total time a job spends in the system from its arrival until the departure
- $\lambda_A \mathbf{E}(R) = \mathbf{E}(X)$ The Little's Law (stationary queuing systems only)

Notes

• A queuing system is *stationary* if the distributions of S_k , W_k , and R_k are independent of k.

Bernoulli single-server queuing process

- Δ frame size
- $p_A = \lambda_A \Delta$ probability of arrival during a frame
- $p_S = \lambda_S \Delta$ probability of service during a frame

Markov property

- $p_{00} = \mathbf{P}\{\text{no arrivals}\} = 1 p_A$
- $p_{01} = \mathbf{P}\{\text{new arrival}\} = p_A$

For all i > 1,

- $p_{i,i-1} = \mathbf{P}\{\text{no arrivals} \cap \text{one departure}\} = (1 p_A)p_S$
- $p_{i,i} = \mathbf{P}\{\text{no arrivals} \cap \text{no departures}\} + \mathbf{P}\{\text{one arrival} \cap \text{one departure}\} = (1 p_A)(1 p_S) + p_A p_S$
- $p_{i,i+1} = \mathbf{P}\{\text{one arrival} \cap \text{no departures}\} = p_A(1 p_S)$

Steady state distribution can be computed if $\lambda_S > \lambda_A$.

M/M/1 system

A/S/n/C

- ullet A distribution of interarrival times
- S distribution of service times
- n number of servers
- C capacity
- $\pi_x = \mathbf{P}\{X = x\} = r^x(1-r) \text{ for } x = 0, 1, 2, \dots$
- $\mathbf{E}\{X\} = \frac{r}{1-r}$
- $Var(x) = \frac{r}{(1-r)^2}$
- $r = \lambda_A/\lambda_S = \mu_S/\mu_A = 1 \pi_0 = P(busy) = 1 P(idle)$

The system is functional if r < 1. IF $r \ge 1$ the system will be overloaded.

- $W = S_1 + S_2 + S_3 + \ldots + S_X$ waiting time $\mathbf{E}(W) = \mathbf{E}(S_1 + \ldots + S_X) = \mathbf{E}(S)\mathbf{E}(X) = \frac{\mu_S r}{1-r} = \frac{r}{\lambda_S(1-r)}$ Expected
- $\mathbf{E}(R) = \mathbf{E}(W) + \mathbf{E}(S) = \frac{\mu_S r}{1-r} + \mu_S = \frac{\mu_S}{1-r} = \frac{1}{\lambda_S(1-r)}$ Expected response

Queue length (number of waiting jobs)

- $X_w = X X_s$ where X_s is number of jobs getting service at any time (0
- $\mathbf{E}(X_w) = \mathbf{E}(X) \mathbf{E}(X_s) = \frac{r}{1-r} r = \frac{r^2}{1-r}$

Chapter 8 – Introduction to Statistics

- $\sigma(\hat{\theta})$ standard error of estimator $\hat{t}heta$ of parameter θ
- $s(\hat{\theta})$ estimated standard error standard deviation of the estimator

Chapter 9 – Statistical Inference I

Method of moments

- $\begin{array}{ll} \bullet & \mu_k = \mathbf{E}(X^k) \text{k-th population moment} \\ \bullet & m_k = \frac{1}{n} \Sigma_{i=1}^n X_i^k \text{k-th sample moment} \end{array}$

Method of maximum likelihood

- $\mathbf{P}\{\mathbf{X}=(X_1,\ldots,X_n)\}=P(\mathbf{X})=P(X_1,\ldots,X_n)=\prod_{i=1}^n P(X_i)$ discrete case, where P(X) is the pmf of the distribution
- $\mathbf{P}\{x-h < X < x+h\} = \int_{x-h}^{x+h} f(y)dy \approx (2h)f(x)$ continuous case, where f(y) is the density function

Confidence intervals

- $\hat{\theta} \pm z_{\alpha/2} * \sigma(\hat{\theta})$ Confidence interval, Normal distribution (1α) 100% confidence interval for θ .
- $\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ Confidence interval for the mean; σ is known
- $\bar{X} \bar{Y} \pm z_{\alpha/2} \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}$ Confidence interval for the difference of means; known σ 's
- $n \ge (z_{\alpha/2}\sigma/\Delta)^2$ sample size for significance, where Δ is margin of error

Unknown standard deviation

- $\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ Confidence interval for a population proportion
- $\hat{p}_1 \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ Confidence interval for the difference of proportions
- $\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$ Confidence interval for the mean; σ is unknown where $t_{\alpha/2}$ is a critical value from T-distribution