STAT 4033

Exam 2 Notes Name: Cooper Morris

C2S4:

Random Variable: X, assigns a numeric value to outcomes in a sample space.
Discrete Random Variable: Possible values are countable, sum of two dice.
Continuous Random Variable: Possible values continuous, weight of random person.
Probability Mass Function: $P(X = x) = p(x)$
Cumulative Distribution Function: $P(X \leq x)$
Denoted by $F(x)$
$0 \leq p(x) \leq 1$
$\sum_x p(x) = 1$
Support: the set of values such that $f(x) > 0$
Percentile: x_p such that $P(X \leq x_p) = \frac{p}{100}$
$F(x_p) = \int_{-\infty}^{x_p} f(x)\, dx = \frac{p}{100}$

C2S6:

Jointly Distributed Random Variable: Two or more random variables that are related when considering "individuals" in a population.
Joint Probability Mass Function: $P(X = x, Y = y) = p(x, y)$ $P(X = x \cap Y = y)$
Marginal Probability Mass Function:
$P_x(x) = \sum_y p(x, y)$
$f_x(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$
Summing or integrating out the opposite variable gives you the marginal probability mass function for the variable you desire.

C4S1: The Bernoulli Distribution

X $\sim$ Bernoulli(P)
P is the probability of a success happening.
$P(X = x) = p^X(1 - p^{1-x})$
$\mathbb{E}[X] = \mu_X = p$
$\text{Var}(X) = p \cdot (1 - p)$

C4S2: The Binomial Distribution

Chaining together n independent Bernoulli trials. Think of a for loop running n times. $X \sim \text{Bin}(n, p)$
$p(x) = \binom{n}{x} p^x (1 - p)^{n-x}$
$\mathbb{E}[X] = \mu_X = n \cdot p$
$\text{Var}(X) = n \cdot p \cdot (1 - p)$
Mass function for $X \sim \text{Bin}(n, p)$:
$P(x) = P(X = x) = \binom{n}{x} \cdot p^x \cdot (1 - p)^{n-x}$
Table A1 gives probabilities for the Binomial distribution in the form $P(X \leq x)$
$\mathbb{E}[X] = \mu_X = n \cdot p$
$\sigma_X^2 = \text{Var}(X) = n \cdot p \cdot (1 - p)$
$\sigma_X = \sqrt{n \cdot p \cdot (1 - p)}$
$\hat{p} = \frac{x}{n} = \frac{\sum Y_i}{n}$ where $\hat{p}$ is an estimator of p.
$\mathbb{E}[\hat{p}] = \mu_{\hat{p}} = p$
$\text{Var}(\hat{p}) = \frac{p \cdot (1-p)}{n}$

C4S3: The Poisson Distribution

Taking the limit of the binomial distribution
$\lambda = n \cdot p$
$X \sim \text{Poisson}(\lambda)$
$p(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$
$\mathbb{E}[X] = \mu_X = \lambda$
$\text{Var}(X) = \lambda$
Estimating over t units over time or space:
$\hat{\lambda} = \frac{x}{t}$
$\mathbb{E}[\hat{\lambda}] = \lambda$
$\text{Var}(\hat{\lambda}) = \frac{\lambda}{t}$

C4S4: The Hypergeometric Distribution

A finite population of size N, divided into two groups. R items in group one and N-R in group two. n items selected without replacement.
$X \sim H(N, R, n)$
Probability Mass Function:
$p(x) = \frac{\binom{R}{X}\binom{N-R}{n-x}}{\binom{N}{n}}$
$\mathbb{E}[X] = \mu_X = n \cdot \frac{R}{N}$
$\text{Var}(X) = n\frac{R}{N} \cdot (1 - \frac{R}{N}) \cdot (\frac{N-n}{N-1})$

Geometric Distribution:

A sequence of independent Bernoulli trials where p is constant and X is the number of trials up to and including the first success. Think of a while loop with the first success as the terminating condition.
$X \sim \text{Geom}(P)$
Probability Mass Function:
$p(x) = (1 - p)^{x-1} \cdot p$
$\mathbb{E}[X] = \mu_X = \frac{1}{p}$
$\text{Var}(X) = \frac{1-p}{p^2}$
$\sigma_X = \sqrt{\frac{1-p}{p^2}}$

Discrete Random Variables:
$\mu_x = \mathbb{E}[X] = \sum_x x \cdot p(x)$
$\sigma^2 = \sum_x x^2 \cdot p(x) - \mu_x^2$
$\mathbb{E}[g(X)] = \sum_x g(x) \cdot p(x)$
Continuous Random Variables:
$P(a \leq X \leq b) = \int_a^b f(x)\,dx$
$\mu_x = \mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$
$\sigma^2 = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx - \mu_x^2$
$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f(x)\,dx$

$(\mu \pm k\sigma_x)$
$P(|X - \mu_x| > k\sigma_x) = \frac{1}{k^2}$
C2S5:
Linear Functions of Random Variables
Linear Function: $Y = aX + b$
Expected Value: $a\mathbb{E}[X] + b$
Variance: $a^2 \cdot \text{Var}(X)$
Standard Deviation: $|a| \cdot \sigma_x$

Linear Combinations of Random Variables
Linear Combinations: $\sum_{i=1}^n c_i X_i$
Expected Value: $\sum_{i=1}^n c_i \mathbb{E}[X_i]$
Variance: $\sum_{i=1}^n c_i^2 \text{Var}(X)$ where X_is are independent
Var(x+y) = Var(x) + Var(y)
Var(x-y) = Var(x) + Var(y)

Simple Random Samples
$\bar{x} = \sum_{i=1}^n \frac{1}{n} \cdot x_i$
$\mu_{\bar{x}} = \mathbb{E}[\bar{x}] = \mu$
$\sigma_{\bar{x}}^2 = \text{Var}(\bar{x})$
$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$
C2S6:
$0 \leq p(x,y) \leq 1$
$\sum_x \sum_y p(x,y) = 1$
$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$
$P(a \leq x \leq b, c \leq y \leq d) = \int_c^d \int_a^b f(x,y)\,dx\,dy$
$\mathbb{E}[h(X,Y)] = \sum_x \sum_y h(x,y)p(x,y)$
$\mathbb{E}[h(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y)f(x,y)\,dx\,dy$
$\text{Cov}(X,Y) = \mu_{XY} - \mu_x\mu_y$
$\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$
C3:
$\sigma_{g(X)} \approx |g'(X)| \cdot \sigma_X$
C4S1: The Bernoulli Distribution
$P(X = x) = p^X(1 - p^{1-x})$
$\mathbb{E}[X] = \mu_X = p$
$\text{Var}(X) = p \cdot (1 - p)$

C4S2: The Binomial Distribution
$X \sim \text{Bin}(n,p)$
$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$
$\mathbb{E}[X] = \mu_X = n \cdot p$
$\text{Var}(X) = n \cdot p \cdot (1 - p)$
$P(x) = P(X = x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$
$\mathbb{E}[X] = \mu_X = n \cdot p$
$\sigma_X^2 = \text{Var}(X) = n \cdot p \cdot (1-p)$
$\sigma_X = \sqrt{n \cdot p \cdot (1-p)}$
$\hat{p} = \frac{x}{n} = \frac{\sum Y_i}{n}$ where $\hat{p}$ is an estimator of p.
$\mathbb{E}[\hat{p}] = \mu_{\hat{p}} = p$
$\text{Var}(\hat{p}) = \frac{p \cdot (1-p)}{n}$
C4S3: The Poisson Distribution
$\lambda = n \cdot p$
$X \sim \text{Poisson}(\lambda)$
$p(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$
$\mathbb{E}[X] = \mu_X = \lambda$
$\text{Var}(X) = \lambda$
C4S4: The Hypergeometric Distribution
$X \sim H(N,R,n)$
$p(x) = \frac{\binom{R}{X}\binom{N-R}{n-x}}{\binom{N}{n}}$
$\mathbb{E}[X] = \mu_X = n \cdot \frac{R}{N}$
$\text{Var}(X) = n\frac{R}{N} \cdot (1 - \frac{R}{N}) \cdot (\frac{N-n}{N-1})$
Geometric Distribution:
$X \sim \text{Geom}(P)$
$p(x) = (1-p)^{x-1} \cdot p$
$\mathbb{E}[X] = \mu_X = \frac{1}{p}$
$\text{Var}(X) = \frac{1-p}{p^2}$
$\sigma_X = \sqrt{\frac{1-p}{p^2}}$