Math 480 Exam 3 is Wednesday, Dec. 2. **You are allowed 11 sheets of notes and a calculator.** The first page of exam 1 review is useful. The exam emphasizes HW8-11, and Q9-11. The **final** is Friday, December 11, 2:45-4:45.

From Exam 2 review, know iterated expectations and the conditional variance formula 48), 49), and the Stochastic Processes material 50-65).

Markov Chains ch. 4

66) A (finite or finite state) *Markov chain* $\{X_n : n = 0, 1, 2, ...\}$ is a discrete stochastic process for which time only takes on integer values. X_n will have J possible values 1, ..., J called states. If $X_n = i \in \{1, ..., J\}$, then the Markov chain is in state i at time n. Suppose $x_k \in \{1, ..., J\}$ for $k \geq 0$. The *Markov property* is

$$P(X_{n+1} = j | X_n = i, X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, ..., X_1 = x_1, X_0 = x_0) = P(X_{n+1} = j | X_n = i)$$

for any $n \geq 1$. Hence the conditional probability of X_{n+1} given the past only depends on the state the Markov chain is in at time X_n. Or, given $X_n = i$, then X_{n+1} is independent of the rest of the past (time periods 0, 1, ..., $n-1$). If $0 \leq d < n$ then $P(X_{n+1} = j | X_n = i, X_{n-1} = x_{n-1}, ..., X_d = x_d) = P(X_{n+1} = j | X_n = i)$.

67) **Know:** The *transition probability* $p_{ij} = P(X_{n+1} = j | X_n = i)$. The *transition probability matrix*

$$\boldsymbol{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1J} \\ p_{21} & p_{22} & \cdots & p_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ p_{J1} & p_{J2} & \cdots & p_{JJ} \end{bmatrix}.$$

68) The sum of the probabilities in any row of $\boldsymbol{P}$ is $\sum_{j=1}^{J} p_{ij} = 1$ for row $i = 1, ..., J$.

69) For small J, a transition diagram list the J states with J arrows leaving each state and J arrows entering each state. Then there are J^2 arrows corresponding to the p_{ij} that form $\boldsymbol{P}$. An arrow labelled p_{ij} goes from state i to state j. An arrow labelled p_{ii} goes from state i to state i. A variant on the transition diagram leaves out p_{ii}, which can be found using 68), and leaves out any arrow corresponding to $p_{ij} = 0$ for $i \neq j$.

70) **Know:** $P(X_{n+m} = j | X_m = i) = p^n_{ij}$ where p^n_{ij} is the ijth entry of $\boldsymbol{P}^n = \boldsymbol{P}\boldsymbol{P} \cdots \boldsymbol{P}$ where there are n matrices $\boldsymbol{P}$ in the multiplication. This formula is for a homogeneous Markov chain where the transition probability matrix does not depend on the time period j, so $\boldsymbol{P} = \boldsymbol{P}^{(j)}$ for $j = 0, 1, 2, ...$

71) State j is accessible from state i if $p^n_{ij} > 0$ for some $n \geq 0$. Then, starting in state i, it is possible that the process will enter state j in a finite number of steps.

72) Two states i and j that are accessible to each other *communicate*, written $i \leftrightarrow j$.

73) States that communicate with each other form an equivalence class. A Markov chain is *irreducible* if there is only one class, so all states communicate.

74) For state i, let r_i denote the probability, starting in state i, that the process will ever reenter state i. State i is recurrent if $r_i = 1$ and transient if $r_i < 1$. State i is absorbing if $p_{ii} = 1$ so that the other entries in the ith row are 0. Once in an absorbing

state, such as death, the Markov chain stays in the absorbing state. An absorbing state is recurrent. All of the states in an irreducible Markov chain are recurrent.

75) A recurrent state will be visited infinitely often. A transient state is not certain to be revisited and will only be visited a finite number of times. Hence a Markov chain must have at least on recurrent state to run indefinitely for $n = 1, 2, ...$ Starting in a transient state i, the number of time periods N the process will be in state i, including the initial time, is geometric with finite mean $E(N) = 1/(1 - r_i)$. State i is recurrent if $E(N) = \infty$ and is transient if $E(N) < \infty$.

76) If state i is recurrent and $i \leftrightarrow j$, then state j is recurrent. If state i is transient and $i \leftrightarrow j$, then state j is transient.

77) **Know:** Let $\boldsymbol{\pi}_n = (\pi_{1n}, ..., \pi_{Jn})$ denote the vector of probabilities of being in states 1 to J at time n. Let $\boldsymbol{\pi}_0 = (\pi_{10}, ..., \pi_{J0})$ where $\pi_{i0} = P(X_0 = i)$ is the probability that the process is in state i at the start, time 0. Then $\boldsymbol{\pi}_n$ is the *state vector* at time n and

$$\boldsymbol{\pi}_n = \boldsymbol{\pi}_0 \boldsymbol{P}^n = \boldsymbol{\pi}_1 \boldsymbol{P}^{n-1} = \boldsymbol{\pi}_2 \boldsymbol{P}^{n-2} = \cdots = \boldsymbol{\pi}_k \boldsymbol{P}^{n-k} = \cdots = \boldsymbol{\pi}_{n-1} \boldsymbol{P}$$

and $\boldsymbol{\pi}_{n+1} = \boldsymbol{\pi}_n \boldsymbol{P}$. This formula is for a homogeneous Markov chain.

78) $\boldsymbol{\pi}_0$ is the initial distribution of the Markov chain. Either $\boldsymbol{\pi}_0$ is given or the problem states that the Markov chain starts in state j. Then $\boldsymbol{\pi}_0 = (0, ..., 0, 1, 0, ..., 0)$ where the 1 is in position j.

79) For a homogeneous Markov chain, we could have $\boldsymbol{P}^n \to \boldsymbol{P}^\infty$ as $n \to \infty$, or we could have $\boldsymbol{P}^n$ periodic: $\boldsymbol{P}^n$ takes on K matrices as $n \to \infty$ and does not converge. A homogeneous irreducible Markov chain that is not periodic is *aperiodic*, and has $\boldsymbol{P}^n \to \boldsymbol{P}^\infty$ and $\lim_{n\to\infty} \boldsymbol{\pi}_n = \boldsymbol{\pi}_\infty = \boldsymbol{\pi} = (\pi_1 \ \cdots \ \pi_J)$ where $\boldsymbol{\pi}$ is the *stationary distribution* of the irreducible aperiodic Markov chain. Here $\pi_j = \pi_{j\infty}$ is the long run proportion of time the chain is in state j for $j = 1, ..., J$.

80) An irreducible aperiodic Markov chain has $\lim_n P_{ij}^n = \pi_j$ as $n \to \infty$ so

$$\boldsymbol{P}^n \to \boldsymbol{P}^\infty = \begin{bmatrix} \boldsymbol{\pi} \\ \boldsymbol{\pi} \\ \vdots \\ \boldsymbol{\pi} \end{bmatrix}.$$

Hence each row of $\boldsymbol{P}^\infty = \boldsymbol{\pi}$. Note that $\boldsymbol{\pi}\boldsymbol{P}^\infty = \boldsymbol{\pi}$. It is also true that $\boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi}$.

81) **Know:** For a *nonhomogeneous Markov chain*, the matrix of transition probabilities $\boldsymbol{P}^{(k)}$ depends on the kth step of the process. Then $\boldsymbol{\pi}_n$ = state vector at time n satisfies $\boldsymbol{\pi}_n = \boldsymbol{\pi}_0 \boldsymbol{P}^{(1)} \boldsymbol{P}^{(2)} \cdots \boldsymbol{P}^{(n)}$.

Sometimes the following notation is used $\boldsymbol{P}^{(j)} = \boldsymbol{P}_j = \boldsymbol{Q}^{(j)} = \boldsymbol{Q}_j$.

82) **Know:** For hand calculations multiply the state vector times the matrix. Avoid multiplying matrices. So $\boldsymbol{\pi}_3 = (\boldsymbol{\pi}_0 \boldsymbol{P}^{(1)}) \boldsymbol{P}^{(2)} \boldsymbol{P}^{(3)} = (\boldsymbol{\pi}_1 \boldsymbol{P}^{(2)}) \boldsymbol{P}^{(3)} = \boldsymbol{\pi}_2 \boldsymbol{P}^{(3)}$ for a nonhomogeneous Markov chain, and $\boldsymbol{\pi}_3 = (\boldsymbol{\pi}_0 \boldsymbol{P}) \boldsymbol{P} \boldsymbol{P} = (\boldsymbol{\pi}_1 \boldsymbol{P}) \boldsymbol{P} = \boldsymbol{\pi}_2 \boldsymbol{P}$ for a homogeneous Markov chain.

ch. 10 Brownian Motion

83) A stochastic process $\{X(t), t \geq 0\}$ is a *Brownian motion* or Wiener process if i) $X(0) = 0$, ii) the process has independent and stationary increments, iii) for every $t > 0$, $X(t) \sim N(0, \sigma^2 t)$. Hence $E[X(t)] = 0$ and $V[X(t)] = \sigma^2 t$ for all $t \geq 0$. iv) $X(t)$ is continuous but nowhere differentiable in t.

See Exam 2 review 54) and 55) for independent and stationary increments.

84) Let $0 \leq t_1 < t_2$. Then $X(t_2) - X(t_1) \sim N[0, \sigma^2(t_2 - t_1)]$. Let $t, s > 0$. Then $X(t+s) - X(s) \sim N(0, \sigma^2 t)$.

85) **Know**: When $\sigma = 1$, the process in 83) is called a standard Brownian motion $\{Z(t), t \geq 0\}$. For $s, t > 0$, $Z(t) - Z(s) \sim N(0, |t - s|)$.

86) $W = Z(t+s)|Z(t) \sim N[Z(t), s]$ with $E(W) = Z(t)$ and $V(W) = s$. For a story problem, $t + s - t = s$. Hence if $W = Z(12)|Z(3) = 52$, then $t = 3$ and $s = 12 - 3 = 9$. Thus $W \sim N(52, 9)$.

Note that $Z(t+s) = Z(t+s) - Z(t) + Z(t)$ and $Z(t+s) - Z(t) \sim N(0, s) \perp\!\!\!\perp Z(t)$. Given $Z(t)$, the value $Z(t)$ acts as a constant.

87) Let $s < t$. Then $W \sim Z(s)|Z(t) = B \sim N\left[\frac{sB}{t}, \frac{s}{t}(t-s)\right]$ with $E(W) = sB/t$ and $V(W) = s(t-s)/t$.

88) A stochastic process $\{B(t), t \geq 0\}$ is a Brownian motion with drift coefficient μ and variance parameter σ^2 if $B(t) = \mu t + \sigma Z(t)$ where $Z(t)$ is a standard Brownian motion. Then i) $B(0) = 0$, ii) $B(t)$ has stationary and independent increments, iii) $B(t) \sim N(\mu t, \sigma^2 t)$. The volatility of the process is σ.

89) By 90), $B(t)$ is an arithmetic Brownian motion with $B(0) = 0$. Hence $B(t+s) - B(t) \sim N(\mu s, \sigma^2 s)$, and $B(t+s)|B(t) \sim N(B(t) + \mu s, \sigma^2 s)$ where $s, t > 0$.

90) A stochastic process $\{A(t), t \geq 0\}$ is an arithmetic Brownian motion with drift coefficient μ and variance parameter σ^2 if $A(t) = A(0) + \mu t + \sigma Z(t)$ where $Z(t)$ is a standard Brownian motion. Then $A(t) \sim N(A(0) + \mu t, \sigma^2 t)$, $A(t) - A(0) \sim N(\mu t, \sigma^2 t)$, $A(t+s) - A(t) \sim N(\mu s, \sigma^2 s)$, and $A(t+s)|A(t) \sim N(A(t) + \mu s, \sigma^2 s)$ where $s, t > 0$.

91) A stochastic process $\{G(t), t \geq 0\}$ is a geometric Brownian motion if $\log(G(t)) = A(t)$. See 90). Then $G(t) \sim \text{lognormal}(A(0)+\mu t, \sigma^2 t)$, and $G(t)/G(0) \sim \text{lognormal}(\mu t, \sigma^2 t)$. By 92), $E(G(t)) = \exp(A(0) + \mu t, 0.5\sigma^2 t)$ and $E[G(t)/G(0)] = \exp(\mu t, 0.5\sigma^2 t)$.

92) If $Y \sim N(\mu, \sigma^2)$, then $X = e^Y \sim \text{lognormal}(\mu, \sigma^2)$. Then $E(X^j) = E(e^{jY}) = \phi_Y(j)$ where $\phi_Y(t) = \exp(\mu t + 0.5\sigma^2 t)$.

Ch. 11: Simulation

93) Here are some distributions with the pdf $f(x)$, cdf $F(x)$, and $F^{-1}(u)$. Recall that you solve $u = F(x)$ for $x = F^{-1}(u)$ where $0 < u < 1$.

a) Exponential(λ)= Gamma($\alpha = 1, \lambda$): $f(x) = \lambda e^{-\lambda x}$ where x, $\lambda > 0$. $F(x) = 1 - e^{-\lambda x}$, $F^{-1}(u) = -\frac{1}{\lambda}\log(1-u)$.

b) (two parameter) Pareto(α, θ): $f(x) = \frac{\alpha\theta^\alpha}{(\theta + x)^{\alpha+1}}$ where α, θ, and x are positive. $F(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^\alpha$, $F^{-1}(u) = \theta[(1-u)^{-1/\alpha} - 1]$.

c) If $X \sim$ single parameter Pareto(α, θ): $f(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}} I(x > \theta)$ where $\alpha > 0$ and θ is real. Note the **support** is $x > \theta$. $F(x) = 1 - \left(\frac{\theta}{x}\right)^{\alpha}$ for $x > \theta$. $F^{-1}(u) = \theta[(1-u)^{-1/\alpha}]$.

d) Weibull(θ, τ): $f(x) = \dfrac{\tau (x/\theta)^{\tau} e^{-(x/\theta)^{\tau}}}{x}$ where $\theta > 0$ and $\tau > 0$.
$F(x) = 1 - e^{-(x/\theta)^{\tau}}$, $F^{-1}(u) = \theta[-\log(1-u)]^{1/\tau}$.
If $X \sim EXP(\lambda)$, then $X \sim$ Weibull$(\theta = 1/\lambda, \tau = 1)$.

e) Inverse Weibull(θ, τ): $f(x) = \dfrac{\tau (\theta/x)^{\tau} e^{-(\theta/x)^{\tau}}}{x}$.
Here $\theta, \tau > 0$ and the Inverse Weibull$(\theta, \tau = 1)$ RV is the Inverse Exponential(θ) RV. $F^{-1}(u) = \theta[-\log(u)]^{-1/\tau}$.

f) If $X \sim N(\mu, \sigma^2)$, then the cdf $F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$, $Z \sim N(0,1)$. $F^{-1}(u) = \mu + \sigma z_u$ where $\sigma > 0$, μ is real, and $P(Z \le z_u) = u$ with $Z \sim N(0,1)$.

g) If $X \sim$ lognormal(μ, σ^2): $F(x) = \Phi\left(\frac{\log(x) - \mu}{\sigma}\right)$, If $X \sim LN(\mu, \sigma)$, then $\log(X) \sim N(\mu, \sigma^2)$. Here $\sigma > 0$ and μ is real. $F^{-1}(u) = \exp(\mu + z_u \sigma)$.

94) **Inversion Method for a pdf** $X_i = F^{-1}(U_i)$: Let X be from a distribution with increasing cdf $F(x)$. Let $u_1, ..., u_n$ be pseudo U(0,1) random numbers. Then $x_1 = F^{-1}(u_1), ..., x_n = F^{-1}(u_n)$ are pseudo random numbers from the distribution of X. So $x_i = F^{-1}(u_i)$ where $F^{-1}(u)$ is given for several brand name distributions in 93). If $F^{-1}(u)$ is not given, solve $u = F(x)$ for $x = F^{-1}(u)$ and use $x_i = F^{-1}(u_i)$. Sometimes need to get the cdf $F(X) = \int_0^x f(t)dt$ where $f(t)$ is the pdf of a RV X with support $x > 0$.

95) Normal approximation to the binomial. Let Y count the number of successes in n trials where the probability of a success in p then Y is binomial$(n = 50, p = 0.3)$. Let X be a normal RV with mean $\mu = np$ and SD $\sigma = \sqrt{np(1-p)}$. Then $P(Y \ge 18) = P(X \ge 17.5)$ and $P(Y \le 18) = P(X \le 18.5)$. Ideally, this approximation should not be used unless $n > 9p/(1-p)$ and $n > 9(1-p)/p$. In general, replace integer 18 by integer i.

96) Normal approximation to the Poisson. Let $Y \sim$ Poisson(λ) where $\lambda \ge 9$. Let Y be a normal RV with mean $\mu = \lambda$ and variance $\sigma^2 = \lambda$. Then $P(Y \ge 18) = P(X \ge 17.5)$ and $P(Y \le 18) = P(X \le 18.5)$. In general, replace integer 18 by integer i.

97) If the data $X_1, ..., X_n$ is arranged in ascending order from smallest to largest and written as $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$, then $X_{(i)}$ is the ith order statistic and the $X_{(i)}$'s are called the *order statistics*. Let the X_i be iid with cdf F. Let x_α be the 100 αth percentile (quantile) of X: $F^{-1}(\alpha) = x_\alpha$ and $F(x_\alpha) = \alpha$ where $0 < \alpha < 1$. Let the greatest integer function $\lfloor x \rfloor =$ the greatest integer $\le x$, i.e. $\lfloor 7.7 \rfloor = 7$. Let $j = \lfloor (n+1)\alpha \rfloor$. The sample percentile $\hat{x}_\alpha = (1-h)X_{(j)} + hX_{(j+1)}$ for some h where $0 \le h \le 1$. For hand calculations, take $h = 1$ so $\hat{x}_\alpha = X_{(j+1)}$ (take $\hat{x}_\alpha = X_{(n)}$ if $j = n$).

98) Consider intervals that contain c cases $[X_{(1)}, X_{(c)}], [X_{(2)}, X_{(c+1)}], ..., [X_{(n-c+1)}, X_{(n)}]$. Compute $X_{(c)} - X_{(1)}, X_{(c+1)} - X_{(2)}, ..., X_{(n)} - X_{(n-c+1)}$. Then the estimator shorth(c) $= [X_{(s)}, X_{(s+c-1)}]$ is the interval with the shortest length. A large sample $100(1-\delta)\%$ prediction interval (PI) (L_n, U_n) is $P(X_f \in (L_n, U_n)) \to \tau \ge 1 - \delta$ as $n \to \infty$. The shorth(c) interval is a large sample $100(1-\delta)\%$ PI if $c/n \to 1 - \delta$ as $n \to \infty$ that often has the asymptotically shortest length.

99) **Inversion Method for a pmf:** Suppose the pmf has support $0, 1, ..., d, ..., J$ where $J = \infty$ is possible. Let $u_{(n)}$ be the largest U(0,1) pseudo number where $F(d-1) <$

$u_{(n)} \leq F(d)$. Given $u_1, ..., u_n$, set $x_i = j$ if $F(j-1) \leq u_i < F(j)$ where $F(-1) = 0$, so set $x_i = 0$ if $0 \leq u_i < F(0)$. See the table below.

k	$p_k = P(X = k)$	$F(k)$	range of u	resulting x_i
0	p_0	$p_0 = F(0)$	$0 \leq u < F(0)$	0
1	p_1	$p_1 + F(0) = F(1)$	$F(0) \leq u < F(1)$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
j	p_j	$p_j + F(j-1) = F(j)$	$F(j-1) \leq u < F(j)$	j
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
d	p_d	$p_d + F(d-1) = F(d)$	$F(d-1) \leq u < F(d)$	d

100) One way to generate U(0,1) pseudo RVs is to use $x_{n+1} = (a * x_n + c) \ \%\% \ m$ for $n \geq 0$ where $x_0 = d$ is a seed. Then take $u_i = x_i/m$ for $i = 1, ..., k$. The R modulo function %% computes the remainder: $e_1 \ \%\% \ e_2 = e_1 - (\lfloor e_1/e_2 \rfloor)e_2$.) For example, use $x_{n+1} = (69069 * x_n + 1) \ \%\% \ (2^{32})$ and take $u_i = x_i/2^{32}$ for $i = 1, ..., k$ where the seed $x_0 = d$, e.g. $d = 12345$. The R program `runif` is better.

ch. 7: Renewal Theory:

101) Let $\{N(t), t \geq 0\}$ be a counting process, and let X_n be the interarrival time or waiting time between the $(n-1)$th and nth events counted by the process, $n \geq 1$. If the nonnegative X_i are iid, then $\{N(t), t \geq 0\}$ is a *renewal process*.

102) A Poisson process with rate λ is a renewal process where the X_i are iid EXP(λ).

103) As with the Poisson process, let $S_n = \sum_{i=1}^n X_i$ = the time of occurrence of the nth event = waiting time until the nth event. Note that $S_n = S_{n-1} + X_n$ is a random walk with $S_n = Y_n$, $Y_0 = y_0 = 0$ and $e_i = X_i$. Let $E(X_i) = \mu > 0$. Then $E(S_n) = n\mu$ and $V(S_n) = nV(X_i)$ if $V(X_i)$ exists.

104) If $E(X_i) = \mu = \infty$, take $1/\mu = 0$. Then $\dfrac{N(t)}{t} \to \dfrac{1}{E(X_i)} = \dfrac{1}{\mu}$ as $t \to \infty$. When an event occurs we say a renewal has taken place. Then $1/\mu$ is the rate of the renewal process. Since the average time between renewals is μ, the average rate of renewal is 1 every μ time units.

105) If the X_i are the iid waiting times of a renewal process, then

$$E[X_1 + \cdots + X_{N(t)+1}] = E[\sum_{i=1}^{N(t)+1} X_i] = E(X)E[N(t)+1].$$

106) Let $m(t) = E[N(t)]$ be the renewal function = mean value function. *Elementary Renewal Theorem:* $\dfrac{m(t)}{t} = \dfrac{E[(N(t)]}{t} \to \dfrac{1}{E(X_i)} = \dfrac{1}{\mu}$ as $t \to \infty$.

107) Normal approximation (CLT for a renewal process): Let $E(X_i) = \mu$ and $V(X_i) = \sigma^2$. Then

$$\frac{N(t) - \frac{t}{\mu}}{\sqrt{\frac{t\sigma^2}{\mu^3}}} \xrightarrow{D} N(0,1)$$

as $t \to \infty$. Thus

$$N(t) \approx N\left(\frac{t}{\mu}, \frac{t\sigma^2}{\mu^3}\right).$$

108) Let $W_1, W_2, \ldots$ be iid Bernoulli$(p) \sim bin(1,p)$ random variables. Let the Bernoulli process $N(t)$ count the number of 1's = number of "successes." Then the waiting times X_i are iid geom(p), and $S_n \sim NB(n,p)$. Then $N(t) = \sum_{i=1}^{\lfloor t \rfloor} W_i \sim bin(\lfloor t \rfloor, p)$ for $t \geq 1$, and $N(t) = 0$ for $0 \leq t < 1$.

109) If Y has a *hypergeometric distribution*, $Y \sim \mathrm{HG(C, N-C, n)}$, then the data set contains N objects of two types. There are C objects of the first type (that you wish to count or "success") and $N - C$ objects of the second type. Suppose that n objects are selected at random without replacement from the N objects. Then Y counts the number of the n selected objects that were of the first type. The pmf of Y is

$$p(y) = P(Y = y) = \frac{\binom{C}{y}\binom{N-C}{n-y}}{\binom{N}{n}}$$

where the integer y satisfies $\max(0, n - N + C) \leq y \leq \min(n, C)$. The right inequality is true since if n objects are selected, then the number of objects y of the first type must be less than or equal to both n and C. The first inequality holds since $n - y$ counts the number of objects of second type. Hence $n - y \leq N - C$.

Let $p = C/N$. Then

$$E(Y) = \frac{nC}{N} = np$$

and

$$V(Y) = \frac{nC(N-C)}{N^2}\frac{N-n}{N-1} = np(1-p)\frac{N-n}{N-1}.$$

If n is small compared to both C and $N - C$ then $Y \approx bin(n,p)$. If n is large but n is small compared to both C and $N - C$ then $Y \approx N(np, np(1-p))$.