For extra practice, several additional review problems are printed below. Solutions to these problems can be found on the exams page of the course website. I've tried to sort the problems by whether I think they are more appropriate as an in-class question or as a take-home question. While these questions are representative of the typical scope and difficulty of individual exam questions, this review is not comprehensive, nor does it necessarily represent the total amount of time available for the exam.

In-class Exam.

(1) Surprise eggs are type of collectible popular among children; each egg contains a randomized toy that is revealed once the egg is purchased and opened. Suppose there are a total of n unique toys types, and that each egg has equal chance of containing each type. If you purchase t eggs, what is the expected number of distinct toys you will collect?

(2) Let $Z \sim N(0,1)$ and $Y = |Z|$. We say that Y has the *folded Normal* distribution. Find **two** expressions for the MGF of Y as unsimplified integrals, one integral based on the PDF of Y and one based on the PDF of Z.

(3) Find the probability that the quadratic polynomial $Ax^2 + Bx + 1$ has at least one real root, if A, B are iid $\text{Unif}(0,1)$. What is the expected number of real roots of the polynomial?

(4) A variable X is said to have the *arcsine* distribution if its CDF F is given by

$$F(x) = \frac{2}{\pi}\sin^{-1}(\sqrt{x}) \qquad \text{for } 0 < x < 1$$

and $F(x) = 0$ for $x \le 0$ and $F(x) = 1$ for $x \ge 1$.

(a) Check that F is indeed a valid CDF.

(b) Find the corresponding PDF f. (Recall that the derivative of $\sin^{-1}(x)$ is $\frac{1}{\sqrt{1-x^2}}$)

(c) Find a formula for the quantile function F^{-1} of X.

(d) Suppose $U \sim \text{Unif}(0,1)$ and let F^{-1} be the quantile function of X. What is the name of the distribution of $F^{-1}(U)$?

(5) Let $X \sim \text{Pois}(\lambda)$. (While proofs of each of the following are in the text, it's good practice with Taylor Series to compute them yourself).

(a) Show directly that the PMF for X,

$$p(k) = \frac{e^{-\lambda}\lambda^k}{k!} \quad \text{for} \quad k \ge 0$$

is a valid PMF.

(b) Show using the PMF of X that the mean and variance of X are both λ.

(c) Give an alternative proof of the preceding fact by finding the MGF of X using LOTUS and then computing moments.

(6) Suppose X and Y are independent variables with $X \sim \text{Geom}(p_1)$ and $Y \sim \text{Geom}(p_2)$. Let $q_1 = 1 - p_1$ and $q_2 = 1 - p_2$.

(a) Find $P(X < Y)$. *Hint: recall that the CDF of $Z \sim \text{Geom}(p)$ is $F_Z(z) = 1 - q^{z+1}$ for non-negative integers z.*

(b) What does your answer simplify to in the case when $p_1 = p_2$?

Take-home Exam.

(1) Suppose (X, Y) are random variables with joint PDF given by

$$f(x, y) = c, \ \text{ for } -1 \le x \le 1,\, 0 \le y \le x^2$$

for some constant c.

(a) Sketch the region in the xy-plane where the joint PDF is non-zero.
(b) Find the value of the constant c that makes f a valid joint PDF.
(c) Find the marginal distribution of X, as well as the marginal distribution of Y.
(d) Compute the mean of X and the mean of Y.
(e) Calculate $\mathrm{Cov}(X, Y)$.
(f) Are X and Y independent? Explain.

(2) Suppose $m \ge 2$ and let $X_1, \ldots, X_m$ be iid $\mathrm{Exp}(\lambda)$, let $S_m = \sum_{k=1}^{m} X_k$ and let $M = \min\{X_1, \ldots, X_m\}$. Is S_m exponentially distributed? Is M exponentially distributed? Justify your answer.

(3) Let $A_1, A_2, \ldots, A_m$ be an arbitrary collection of events (we do not assume that the A_i's are either disjoint or independent). Show that

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) \ge \left(\sum_{i=1}^{m} P(A_i) \right) - m + 1$$

by first considering indicator variables and then using the fundamental bridge.

(4) An immortal ant wanders along an infinitely long stick that bears striking resemblance to the x-axis. Suppose at the dawn of time, the ant starts at position 0, and that each second thereafter, the ant moves 1cm in either the positive or the negative direction, each with equal probability, independent of the ant's previous movement. Let X_n be the ant's position at time n.

(a) Compute $E[X_n]$.
(b) Compute $E[|X_n|]$.
(c) Compute $\mathrm{Var}(X_n)$.
(d) Compute $\mathrm{Corr}(X_n, X_{n-1})$ for $n \ge 2$.
(e) Compute $\mathrm{Corr}(X_n, X_1)$ for $n \ge 2$.

(5) A variable X is said to have the *Gumbel Distribution* if $X = -\log Y$ where $Y \sim \mathrm{Expo}(1)$.

(a) Find the CDF of the Gumbel distribution.
(b) Let $X_1, X_2, \ldots$ be iid $\mathrm{Expo}(1)$ and let $M_n = \max\{X_1, \ldots, X_n\}$. Show that $M_n - \log n$ converges in distribution to the Gumbel distribution; that is, show that the CDF of $M_n - \log n$ converges to the Gumbel CDF as $n \to \infty$. *Hint: Recall that* $e^x = (1 + \frac{x}{n})^n$.

(6) Suppose $X \sim \mathrm{Bin}(20, 0.25)$.

(a) Use `dbinom` in R to show that the mean X is 5. *Hint: If `v` and `w` are vectors, then `sum(v*w)` is the sum of pairwise products of entries of `v` and `w`.*
(b) Use `qbinom` in R to show that 5 is a median of X.
(c) Use `dbinom` in R to show that the mode of X is 5. *Note that the function `max(v)` computes the largest value of the vector `v` and that the function `which.max(v)` finds the position of the maximal value of vector `v`.*
(d) Verify that X is **not** symmetric by creating a plot of the PMF in R, despite the fact the mean, median and mode of X are all equal.