#1 on Take-home exam:
cov = 1/80
$E(XY) = \int_0^1 \int_0^y xy6(y-x)dxdy$
$f_{Y|X}(y|x) = \dfrac{f(x,y)}{f_X(x)} = \dfrac{6(y-x)}{3(1-x)^2}$
$E(Y|X=x) = \int_x^1 y f_{Y|X}(y \mid x)dy$
$\{Y \le y\} = \{Y \le y \text{ and } X < x\} \cup \{Y \le y \text{ and } X > x\}$
$F_Y(y) = P(\max\{U_1, U_2, U_3\} \le y) = P(U_1 \le y)^3 = y^3$

For $U(0, 1)$, the parameters are given
For $U(0, \theta)$, $\theta$ is unknown
For $U(\alpha, 1)$, $\alpha$ is unknown
For $U(\alpha, \theta)$, $\alpha$ and $\theta$ are unknown
**Question**: How do we estimate parameters from sample data?
**Example**: {0.2, 0.35, 0.42, 0.49, 0.77}
The data has $U(0, \theta)$ distribution. Then $\theta \ge 0.77$.
e.g. $\theta = 0.77 = \max\{x_1, ..., x_5\}$.
**Example**: $U(\theta, \theta + 1)$, range = 1 unit

**Random Samples and Parametric Families of Distributions**
This is from maybe Ch. 5 or Ch. 4.10.

Consider an infinite population represented by a random variable $X$ with a parametrized pdf or pmf $f(X|\theta)$ where $\theta$ is unknown with a parameter space $\Theta$.
If $n$ observations $X_1, ..., X_n$ are iid from the family of distributions, then $X_1, ..., X_n$ form a random sample.

For finite populations, the idea applies if the sampling is done with replacement.

If sampling is done without replacement, it is called a simple random sampling (in finite populations). In this case, observations are dependent.

For statistical convenience, we can ignore such a dependence if the sample size ($n$) is small compared to the population size ($N$). Usually, a sample size that is less than 5% of the population is sufficient to allow ignoring of the dependence.

**Example**: Consider a population of $N = \{1, 2, ..., 1000\}$. Select 10 numbers from this population randomly without replacement. Define a success to be getting a number larger than 200. Then $p = P(\text{success}) = 800/1000 = 0.8$. There are $\binom{1000}{10}$ different samples.

Compute $P$(10 successes) without replacement.

Let $Q = P(X_1 > 200, X_2 > 200, ..., X_{10} > 200)$

(a) Approximation by assuming independence: ($n/N = 0.01$), the number of successes has a binomial distribution with $n = 10$ and $p = 0.8$.

Thus, $Q = \binom{10}{10} 0.8^{10} \bullet 0.2^{0} \approx 0.107374$.

(b) Exact: Hypergeometric distribution

For population = $N$, sample size = $n$, # of successes in population = $K$, # of successes in sample size = $k$, then

$P(X = x) = \dfrac{\binom{K}{x}\binom{N-K}{n-k}}{\binom{N}{n}}$. So $P(X = 10) = \dfrac{\binom{800}{10}\binom{200}{0}}{\binom{1000}{10}} = 0.106164$, very close to approximation in (a). But, in many cases, the numbers are too large to allow for the computation.

If numbers are too large to even approximate by using binomial distribution, we can use the Poisson distribution.

**Normal Populations and Estimation**

Consider a normal population with parameters μ and $\sigma^2$, $N(\mu, \sigma^2)$, where μ or $\sigma^2$ or both are unknown.

When μ is unknown, a natural estimator from a random sample of n observations is the sample mean $\bar{X} = \dfrac{X_1 + \cdots X_n}{n}$.

When σ is unknown, the most popular estimators for $\sigma^2$ are

(a) $\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ if μ is known

(b) $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ or $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ when μ is unknown.

$S^2$ is unbiased

Facts: $E(\bar{X}) = \mu$, $E(S^2) = \sigma^2$ unbiased estimates.

$\bar{X}$ and $S^2$ are independent.