Probability and stats mop-up unit

- Jointly distributed random variables (spring-break eve)
- Understanding covariance and correlation (Monday)
- Convergence of random variables (**Today**)
- The weak law of large numbers (Friday)
- The central limit theorem (next week Monday)
- Review for test (next week Wednesday)
- Test 2 (next week Friday)

Today:

- Goals: WLLN and CLT
- Markov's and Chebyshev's inequalities
- Limits
- Sequences of random variables
- Kinds of convergence

Theorem (The Weak Law of Large Numbers)

If $X_0, X_1, \ldots X_{n-1}$ are a sequence of independent and identically-distributed random variables, each having mean $E[X_i] = \mu$, then

$$\overline{X}_n \xrightarrow{\mathcal{P}} \mu$$

that is, the sample mean converges in probability to μ.

Theorem (The central limit theorem)

If $X_0, X_1, \ldots X_{n-1}$ are a sequence of independent and identically-distributed random variables, each having mean $E[X_i] = \mu$ and variance σ^2, then

$$\sqrt{n}\,\overline{X}_n \xrightarrow{\mathcal{D}} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

that is, the sample mean (times $\sqrt{n}$) converges in distribution to the normal distribution.

Theorem (Markov's inequality)

If X is a random variable with codomain $\mathbb{R}^{nonneg}$, then for any $a \in \mathbb{R}^+$,

$$P(X \geq a) \leq \frac{E[X]}{a}$$

Theorem (Chebyshev's inequality)

If X is a random variable with mean μ and variance σ^2, then for any $k \in \mathbb{R}^+$,

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

Theorem

If X is a random variable and $Var(X) = 0$, then $P(X = E[X]) = 1$

1. Suppose that the number of items that a certain factory produces is a random variable with mean 50, but we don't know anything else about how that number is distributed.
   a. What can we conclude about the probability that a given week's production will be greater than 75?
   b. Suppose further that we know that the variance of the random variable is 25. What then can we conclude about the probability that a week's production will be between 40 and 60?
2. Let X be a uniformly distributed continuous random variable over the interval $(0, 10)$.
   a. What does Chebyshev's inequality tell us about $P(|X - 5| > 4)$?
   b. What is $P(|X - 5| > 4)$ exactly?

Adapted from Ross, *A First Course in Probability*, 1997 , pg 396–397.

The formal definition of a limit:

Let $f : \mathbb{R} \to \mathbb{R}$ and let $a, L \in \mathbb{R}$. Then the limit of f as x approaches a is L, written as

$$\lim_{x \to a} f(x) = L$$

if for all $\epsilon \in \mathbb{R}^+$, there exists $\delta \in \mathbb{R}^+$ such that for all $x \in \mathbb{R}$, if $0 < |x - a| < \delta$ then $|f(x) - L| < \epsilon$.

The most important aspect of probability theory concerns the behavior of sequences of random variables. . . . The basic question is this: what can we say about the limiting behavior of a sequence of random variables $X_0, X_1, X_2 \ldots$? Since statistics and data mining are all about gathering data, we will naturally be interested in what happens as we gather more and more data.

Wasserman, All of Statistics, 2004

The most important theoretical results in probability theory are limit theorems. Of these the most important are those that are classified either under the heading "laws of large numbers" or under the heading "central limit theorems." Usually, theorems are considered to be laws of large numbers if they are concerned with stating conditions under which the average of a sequence of random variables converges (in some sense) to the expected average. On the other hand, central limit theorems are concerned with determining conditions under which the sum of a large number of random variables has a probability distribution that is approximately normal.

Ross, A First Course in Probability, 1997

Let $X_0, X_1, \ldots$ be a sequence of random variables and let X be another random variable. Let F_n be the cumulative distribution function of X_n and F be the cumulative distribution function of X.

X_n **converges** to X **in probability**, written as $X_n \xrightarrow{\mathcal{P}} X$, if for any $\epsilon \in \mathbb{R}^+$,

$$\lim_{n \to \infty} P(|X_n - X| > \epsilon) = 0$$

X_n **converges** to X **in distribution**, written as $X_n \xrightarrow{\mathcal{D}} X$, if

$$\lim_{n \to \infty} F_n(x) = F(x)$$

X_n **converges** to X **in quadratic mean**, written as $X_n \xrightarrow{\mathcal{QM}} X$, if

$$\lim_{n \to \infty} E[(X_n - X)^2] = 0$$

X_n **converges almost surely** to X, written as $X_n \xrightarrow{\mathcal{AS}} X$, if

$$P(\lim_{n \to \infty} X_n = X) = 1$$

For next time:

Do the correlation exercise (Jupyter notebook) (due Fri, Mar 20)