November 27
revised November 26, 2002, 1128 minutes

In-Class Problems — Week 13, Wed

Problem 1. Let R be the number of heads that come up when we toss n independent coins, where each coin comes up heads with probability p. The random variable R has a binomial distribution. Using the formula for variance of sums, show that $\text{Var}[R] = np(1-p)$.

Problem 2. Provide an example that shows that the variance of the sum of two random variables is not necessarily equal to the sum of their variances, when the random variables are not independent.

Problem 3. The hat-check staff has had a long day, and at the end of the party they decide to return people's hats at random. Suppose that n people have their hats returned at random. We previously showed that the expected number of people who get their own hat back is 1, irrespective of the total number of people. In this problem we will calculate the variance in the number of people who get their hat back.

Let $X_i = 1$ if the ith person gets his or her own hat back and 0 otherwise. Let $S_n = \sum_{i=1}^{n} X_i$, so S_n is the total number of people who get their own hat back. Show that

(a) $\text{E}\left[X_i^2\right] = 1/n$.

(b) $\text{E}[X_i X_j] = 1/n(n-1)$ for $i \neq j$.

(c) $\text{E}\left[S_n^2\right] = 2$. *Hint:* Use (a) and (b).

(d) $\text{Var}[S_n] = 1$.

(e) Explain why you can not use the variance of sums formula to calculate $\text{Var}[S_n]$.

(f) Using Chebyshev's Inequality, show that $\Pr\{S_n \geq 11\} \leq .01$ for any $n \geq 11$.

Problem 4. Prove that $\text{Var}[X+Y+Z] = \text{Var}[X] + \text{Var}[Y] + \text{Var}[Z]$, if X, Y, Z are pairwise independent. Explain why pairwise independence is sufficient.

Reminder: If two random variable are independent, then $\text{E}[X_i X_j] = \text{E}[X_i]\,\text{E}[X_j]$. Given a set of random variables, pairwise independence implies that every possible pair of random variables is independent.

A Appendix

Random variables $R_1, R_2, \ldots$ are *mutually independent* iff

$$\Pr\left\{\bigcap_i [R_i = x_i]\right\} = \prod_i \Pr\{R_i = x_i\},$$

for all $x_1, x_2, \dots \in \mathbb{R}$. They are *k-wise independent* iff $\{R_i \mid i \in J\}$ are mutually independent for all subsets $J \subset \mathbb{N}$ with $|J| = k$.

Theorem (Expectation of a Product). *If $R_1, R_2, \ldots, R_n$ are mutually independent, then*

$$\mathrm{E}[R_1 \cdot R_2 \cdot \cdots \cdot R_n] = \mathrm{E}[R_1] \cdot \mathrm{E}[R_2] \cdots \mathrm{E}[R_n].$$

The *variance*, $\mathrm{Var}[R]$, of a random variable, R, is:

$$\mathrm{Var}[R] ::= \mathrm{E}\left[(R - \mathrm{E}[R])^2\right].$$

Variance can also be equivalently defined as:

$$\mathrm{Var}[R] ::= \mathrm{E}\left[R^2\right] - \mathrm{E}^2[R],$$

Lemma. *For $a, b \in \mathbb{R}$,*

$$\mathrm{Var}[aR + b] = a^2 \,\mathrm{Var}[R]$$

Theorem 4.1. *If $R_1, R_2, \ldots, R_n$ are* pairwise *independent random variables, then*

$$\mathrm{Var}[R_1 + R_2 + \cdots + R_n] = \mathrm{Var}[R_1] + \mathrm{Var}[R_2] + \cdots + \mathrm{Var}[R_n].$$

Theorem (Markov's Theorem). *If R is a nonnegative random variable, then for all $x > 0$*

$$\Pr\{R \geq x\} \leq \frac{\mathrm{E}[R]}{x}.$$

An alternative formulation is

$$\Pr\{R \geq x\,\mathrm{E}[R]\} \leq \frac{1}{x}.$$

Theorem (Chebyshev). *Let R be a random variable, and let x be a positive real number. Then*

$$\Pr\{|R - \mathrm{E}[R]| \geq x\} \leq \frac{\mathrm{Var}[R]}{x^2}.$$

An alternative formulation is

$$\Pr\{|R - \mathrm{E}[R]| \geq x\sigma_R\} \leq \frac{1}{x^2},$$

where $\sigma_R ::= \sqrt{\mathrm{Var}[R]}$ is the standard deviation of R.