In-Class Problems Week 13, Fri.

Problem 1.
A herd of cows is stricken by an outbreak of *cold cow disease*. The disease lowers the normal body temperature of a cow, and a cow will die if its temperature goes below 90 degrees F. The disease epidemic is so intense that it lowered the average temperature of the herd to 85 degrees. Body temperatures as low as 70 degrees, **but no lower**, were actually found in the herd.

(a) Prove that at most 3/4 of the cows could have survived.

Hint: Let T be the temperature of a random cow. Make use of Markov's bound.

(b) Suppose there are 400 cows in the herd. Show that the bound of part (a) is best possible by giving an example set of temperatures for the cows so that the average herd temperature is 85, and with probability 3/4, a randomly chosen cow will have a high enough temperature to survive.

Problem 2.
A gambler plays 120 hands of draw poker, 60 hands of black jack, and 20 hands of stud poker per day. He wins a hand of draw poker with probability 1/6, a hand of black jack with probability 1/2, and a hand of stud poker with probability 1/5.

(a) What is the expected number of hands the gambler wins in a day?

(b) What would the Markov bound be on the probability that the gambler will win at least 108 hands on a given day?

(c) Assume the outcomes of the card games are pairwise independent. What is the variance in the number of hands won per day?

(d) What would the Chebyshev bound be on the probability that the gambler will win at least 108 hands on a given day? You may answer with a numerical expression that is not completely evaluated.

Problem 3.
The proof of the Pairwise Independent Sampling Theorem 21.5.1 was given for a sequence $R_1, R_2, \ldots$ of pairwise independent random variables with the same mean and variance.

The theorem generalizes straighforwardly to sequences of pairwise independent random variables, possibly with *different* distributions, as long as all their variances are bounded by some constant.

Theorem (Generalized Pairwise Independent Sampling). *Let $X_1, X_2, \ldots$ be a sequence of pairwise independent random variables such that* $\text{Var}[X_i] \leq b$ *for some $b \geq 0$ and all $i \geq 1$. Let*

$$A_n ::= \frac{X_1 + X_2 + \cdots + X_n}{n},$$
$$\mu_n ::= \text{E}[A_n].$$

Then for every $\epsilon > 0$,

$$\Pr\{|A_n - \mu_n| > \epsilon\} \leq \frac{b}{\epsilon^2} \cdot \frac{1}{n}. \tag{1}$$

(a) Prove the Generalized Pairwise Independent Sampling Theorem.

(b) Conclude that the following holds:

Corollary (Generalized Weak Law of Large Numbers). *For every $\epsilon > 0$,*

$$\lim_{n\to\infty} \Pr\{|A_n - \mu_n| \leq \epsilon\} = 1.$$

Appendix

Markov's Theorem

If R is a nonnegative random variable, then for all $x > 0$

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

Variance

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

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

[Variance of an indicator variable], I, with $\Pr\{I = 1\} = p$:

$$\text{Var}[I] = p(1-p).$$

[Variance and constants] For constants, a, b,

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

[Variance Additivity] If $R_1, R_2, \ldots, R_n$ are *pairwise* independent variables, then

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

Chebyshev' s Bound

Let R be a random variable, and let x be a positive real number. Then

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

Pairwise Independent Sampling

Theorem. *Let*

$$S_n ::= \sum_{i=1}^{n} G_i,$$
$$A_n ::= \frac{S_n}{n},$$

where $G_1, \ldots, G_n$ are pairwise independent random variables with the same mean, μ, and deviation, σ. Then

$$\Pr\{|A_n - \mu| \geq x\} \leq \frac{1}{n} \cdot \left(\frac{\sigma}{x}\right)^2. \tag{3}$$

Proof. By linearity of expectation,

$$\mathrm{E}[A_n] = \frac{\mathrm{E}\left[\sum_{i=1}^{n} G_i\right]}{n} = \frac{\sum_{i=1}^{n} \mathrm{E}[G_i]}{n} = \frac{n\mu}{n} = \mu.$$

Since the G_i's are pairwise independent, their variances will also add, so

$$\begin{aligned}
\mathrm{Var}[A_n] &= \left(\frac{1}{n}\right)^2 \mathrm{Var}\left[\sum_{i=1}^{n} G_i\right] && \text{(by (2))} \\
&= \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} \mathrm{Var}[G_i] && \text{(additivity)} \\
&= \left(\frac{1}{n}\right)^2 n\sigma^2 \\
&= \frac{\sigma^2}{n}.
\end{aligned}$$

Now letting R be A_n in Chebyshev's Bound yields (3), as required.

■

6.042J / 18.062J Mathematics for Computer Science
Spring 2010