Chapter 7.4: Expected Value and Variance

Monday, August 3

Summary

- Covariance of X and Y is $E([X - E(X)][Y - E(Y)]) = E(XY) - E(X)E(Y)$
- Correlation of X and Y is $\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$.
- $G = (V, E)$. V = set of vertices, E = $\{(u, v), \ldots\}$ = set of edges.
- Digraph: the direction of the edge matters, (u,v) = arc from u to v.
- *degree* of u = number of neighbors, *isolated* = degree 0
- K_n = complete graph on n vertices, C_n = cycle
- *bipartite* if it can be colored with 2 colors so that neighbors have different colors, or if we can write $V = V_1 \cup V_2$ with $V_1 \cap V_2 = \emptyset$ so that every edge in E has one elt. in V_1 and one in V_2.
- Induced subgraph: choose the vertices; the edges are automatic.
- Complement: $\overline{G}$ has all of the edges that G doesn't.

Covariance/Correlation

1. Flip 3 coins. Let X be the number of heads flipped and let Y be the number of times the sequence HH appears. Find $Cov(X, Y)$.

   $Cov(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{8}(2 \cdot 1 + 2 \cdot 1 + 3 \cdot 2) - (3/2)(1/2) = 10/8 - 3/4 = 1/2 > 0$. These variables are (unsurprisingly) positively correlated.

2. (★) Let X be the number of heads in 2 flips of a fair coin, and let Y be the number of tails. Show that the correlation of X and Y is -1.

   $Cov(X, Y) = E(XY) - E(X)E(Y) = \frac{1}{4}(1 \cdot 1 + 1 \cdot 1) - 1 \cdot 1 = -1/2$. Then $Var(X) = Var(Y) = 1/2$, so $Cov(X, Y)/\sqrt{Var(X)Var(Y)} = -1$. See a generalization below.

3. Show that if $X + Y$ is any constant then the correlation of X and Y is -1.

   Say $X = k - Y$, then $Cov(X, Y) = Cov(k - Y, Y) = -Cov(Y, Y) = -Var(Y)$. Then $Var(X) = Var(Y)$, so $Cov(X, Y)/\sqrt{Var(X)Var(Y)} = -Var(Y)/Var(Y) = -1$.

4. Show that the correlation of X and Y is equal to the correlation of $(aX + b)$ and Y.

   This follows from the fact that $Cov(aX + b, Y) = aCov(X, Y)$ and $\sqrt{Var(aX + b)} = a\sqrt{Var(X)}$. Both of these quantities scale up by a and the scaling cancels out.

5. Prove that $Cov(X, Y)^2 \leq Var(X)Var(Y)$ for any random variables X and Y (This is equivalent to the Cauchy-Schwarz inequality. Hint: Find the a that minimizes $Var(X - aY)$ and use the fact that the variance must be at least 0.)

$0 \leq Var(X-aY) = Var(X)+a^2Var(Y)-2aCov(X,Y)$. This is minimized when $a = Cov(X,Y)/Var(Y)$, giving

$$
\begin{aligned}
0 &\leq Var(X) + Cov(X,Y)^2 Var(Y)/Var(Y)^2 - 2\frac{Cov(X,Y)}{Var(Y)}Var(Y) \\
0 &\leq Var(X) - \frac{Cov(X,Y)^2}{Var(Y)} \\
\frac{Cov(X,Y)^2}{Var(Y)} &\leq Var(X) \\
Cov(X,Y)^2 &\leq Var(X)Var(Y)
\end{aligned}
$$

Graphs

1. If $G = (V, E)$ is a graph, show that $|E| \leq \binom{n}{2}$.

   That's how many distinct pairs of vertices there are.

2. ($\bigstar$) Show that $|E(G)| + |E(\overline{G})| = \binom{n}{2}$.

   Any of those possible vertices that aren't in $E(G)$ are in $E(\overline{G})$ and vice versa.

3. The number of graphs with n (labeled) vertices is $2^{\binom{n}{2}}$

   After we fix the n vertices, the number of graphs is equal to the number of subsets of possible edges, which is 2 to the number of edges, which is $2^{\binom{n}{2}}$.

4. ($\bigstar$) Let G have n vertices and m edges. How many induced subgraphs are there? How many spanning subgraphs are there?

   There are 2^n induced subgraphs (all subsets of vertices) and 2^m spanning subgraphs (all subsets of edges).

5. How many spanning subgraphs of K_n are there with exactly m edges?

   $\binom{n}{m}$, since we fix all of the vertices and pick m edges.

6. Show that if G is bipartite then $|E| \leq \lfloor n^2/4 \rfloor$

   $K_{m,o}$ has mo edges in general, and this number is maximized when o and m are as close as possible. If n is even then $(n/2)(n/2) = n^2/4$, and if n is odd then we have $(n+1)/2 \cdot (n-1)/2 = (n^2-1)/4 < n^2/4$.

7. ($\bigstar$) How many edges does K_n have? What about C_n? $K_{m,n}$?

   K_n has $\binom{n}{2}$ edges, C_n has n, $K_{m,n}$ has mn.

8. For which values of n is K_n bipartite? What about C_n?

   K_n is bipartite only when $n \leq 2$. C_n is bipartite precisely when n is even.

9. Describe and count the edges of $\overline{K_n}, \overline{C_n}, \overline{K_{m,n}}$.

   Subtract the number of edges each of these graphs have from $\binom{n}{2}$ to get the number of edges in the complements.

10. ($\bigstar$) Draw a directed graph on the 7 vertices $\{0, 1, \ldots, 6\}$ where (u, v) is an edge if and only if $v \equiv 3u \pmod 7$.

    $0 \to 0$, $1 \to 3 \to 2 \to 6 \to 4 \to 5 \to 1$.

11. Show that in a simple graph with at least two vertices there must be two vertices that have the same degree (Hint: pigeon-hole principle)

    Say the graph has n vertices with $n \geq 2$. The degree number of a vertex can be anywhere from 0 to n-1…n options in all. But if one vertex has degree 0 then no vertex can have degree $n - 1$, because that would require it to share an edge with the degree 0 vertex. There are therefore at least (n-1) vertices with degrees between 1 and (n-2), so by the Pigeonhole principle two of them must have the same degree.

12. ($\bigstar$) How many triangles does the graph K_n contain?

    $\binom{n}{3}$, since any three vertices make a triangle in K_n.

13. (Hard) If G is triangle-free, then $|E(G)| \leq \lfloor n^2/4 \rfloor$ (use inductions in increments of 2, delete both vertices connected by some edge for the inductive step).

    Base case: the inequality holds for $n = 1$ and $n = 2$.

    Inductive step: suppose the inequality holds for n, and consider a graph with $n+2$ vertices. Pick any edge (u, v). Because G is triangle-free, u and v can have no common neighbors (otherwise (u, v, w, u) would make a triangle), so each of the n remaining vertices is connected to at most one of u and v. Therefore if we delete u, v, and all edges connected to either of them, we will have deleted at most $n+1$ edges.

    The remaining graph has n vertices and by inductive hypothesis has at most $n^2/4$ edges, so when we add u and v back in we get that the graph G has at most $\frac{n^2}{4} + (n+1) = \frac{n^2+4n+4}{4} = \frac{(n+2)^2}{4}$ edges. The proof by induction is complete.

Suggested From Rosen

10.2: 1-5, 7, 19, 21-25, 53-55