Math 425, Fall 1997
Prof. Hochster

Solutions to Selected Problems from Assignment #10

Many routine details of integration are omitted. We use the formula $\int e^{-ax} f(x)\,dx = -e^{-ax}(\frac{f(x)}{a} + \frac{f'(x)}{a^2} + \cdots) + c$ when $f(x)$ is a polynomial in x. It is also worth noting that when $g(x)$ is a polynomial in x and $a > 0$, then $-e^{-ax}g(x)\big|_0^\infty$ is the constant term of $g(x)$.

pp. 377–83.

31. (a) Number the non-aces from 1 to 48, and let X_i be 1 if the ith non-ace occurs before the second ace and 0 if not. Then the number of cards turned up until two aces turn up is $X = X_1 + \cdots X_{48} + 2$, where $X_1 + \cdots + X_{48}$ is the number of non-aces and the 2 is the number of aces. Then $E[X] = 48E[X_i] + 2$, and $E[X_i]$ is the probability that a given non-ace occurs before the second ace. Consider only the 5 cards consisting of the given non-ace and the 4 aces. The given non-ace is as likely to hold one position as another among these 5 cards, so the probability that it is first or second is $\frac{2}{5}$. Thus, the answer is $48\frac{2}{5} + 2$. Similarly, (b) is $39\frac{5}{14} + 5$ and (c) is $39\frac{13}{14} + 13$.

32. Let X_i be 1 if the ith roll of the die produces a 1 and 0 otherwise, and let Y_i be 1 if the ith roll of the die produces a 2 and 0 otherwise. Then $X = X_1 + \cdots + X_n$ and $Y = Y_1 + \cdots + Y_n$. Then $\mathrm{Cov}\,(X, Y) = \mathrm{Cov}\,(X_1 + \cdots + X_n, Y_1 + \cdots + Y_n)$, and this is the sum of the n^2 terms $\mathrm{Cov}\,(X_i, Y_j)$ as i and j take on all values between 1 and n. If $i \neq j$ then $\mathrm{Cov}\,(X_i, Y_j) = 0$, since X_i and Y_j are independent in that case, and so the answer is the same as the sum of the n terms $\mathrm{Cov}\,(X_i, Y_i)$, all of which are equal. Each may be computed as $E[X_iY_i] - E[X_i]E[Y_i]$. Now for all i, $X_iY_i = 0$ since $X_i = 0$ unless the ith roll is a 1 and $Y_i = 0$ unless the ith roll is a 2. Thus, $E[X_iY_i] = 0$. Also, both $E[X_i]$ and $E[Y_i]$ are 1/6 (the probability of throwing a 1 in the case of X_i and a 2 in the case of Y_i), so that $\mathrm{Cov}\,(X_i, Y_i) = -1/36$, and so $\mathrm{Cov}\,(X, Y) = -n/36$.

33. Let U be the outcome of the first roll and V of the second. Then $\mathrm{Cov}\,(X, Y) = \mathrm{Cov}\,(U+V, U-V) = \mathrm{Cov}\,(U, U) - \mathrm{Cov}\,(U, V) + \mathrm{Cov}\,(V, U) - \mathrm{Cov}\,(V, V)$. The two terms in the middle are both 0 because U and V are independent, but even if U and V are arbitrary, the middle terms cancel. This gives $\mathrm{Cov}\,(U, U) - \mathrm{Cov}\,(V, V)$. This is 0 without calculation because U and V have the same distribution. [**Note:** $U + V$ and $U - V$ are consequently uncorrelated, but they are **not independent**. For example, if $U + V$ is even, then so is $U - V$. Also, if $U + V = 12$, then $U - V$ must be 0.]

34. To calculate $E[X]$, $E[Y]$, and $E[XY]$ we need to evaluate the iterated integrals $\int_0^\infty \int_0^x x(2e^{-2x}/x)\,dy\,dx$, $\int_0^\infty \int_0^x y(2e^{-2x}/x)\,dy\,dx$, and $\int_0^\infty \int_0^x xy(2e^{-2x}/x)\,dx\,dy$. In the first, the inner integral simplifies to $\int_0^x 2e^{-2x}\,dy = 2e^{-2x}y\big|_0^x = 2xe^{-2x}$ and then we have that $\int_0^\infty 2xe^{-2x}\,dx = -e^{-2x}(\frac{2x}{2} + \frac{2}{2^2})\big|_0^\infty = \frac{2}{4} = \frac{1}{2}$. In the second, the inner integral yields the value $(y^2e^{-2x}/x)\big|_0^x = xe^{-2x}$, and integrating again yields $\frac{1}{4}$. Finally, in the third, the inner integral yields $y^2e^{-2x}\big|_0^x = x^2e^{-2x}$, and $\int_0^\infty x^2e^{-2x}\,dx = -e^{-2x}(\frac{x^2}{2} + \frac{2x}{4} + \frac{2}{8})\big|_0^\infty = \frac{1}{4}$. Then $\mathrm{Cov}\,(X, Y) = E[XY] - E[X]E[Y] = \frac{1}{4} - (\frac{1}{2})(\frac{1}{4}) = \frac{1}{4} - \frac{1}{8} = \frac{1}{8}$.

36. Similarly, here we need to calculate three iterated integrals in order to find $E[X]$, $E[Y]$, and $E[XY]$. We can write the density as $\frac{1}{y}e^{-y}e^{-x/y}$, and the three integrals are $\int_0^\infty \int_0^\infty x\frac{1}{y}e^{-y}e^{-x/y}\,dx\,dy = \int_0^\infty \frac{1}{y}e^{-y}\,(\int_0^\infty xe^{-x/y}\,dx)\,dy$, $\int_0^\infty \int_0^\infty y\frac{1}{y}e^{-y}e^{-x/y}\,dx\,dy = \int_0^\infty e^{-y}\,(\int_0^\infty e^{-x/y}\,dx)\,dy$ and $\int_0^\infty \int_0^\infty xy\frac{1}{y}e^{-y}e^{-x/y}\,dx\,dy = \int_0^\infty e^{-y}\,(\int_0^\infty xe^{-x/y}\,dx)\,dy$. We do the inner integrations treating y as a constant. For the first, $\int_0^\infty xe^{-x/y}\,dx = -e^{-x/y}(\frac{x}{1/y} + \frac{1}{(1/y)^2})\big|_0^\infty = y^2$. Then we have $E[X] = \int_0^\infty \frac{1}{y}e^{-y}(y^2)\,dy = \int_0^\infty ye^{-y}\,dy = -e^{-y}(y+1)\big|_0^\infty = 1$. For the second, the inner integral is $\int_0^\infty e^{-x/y}\,dx = -e^{-x/y}y\big|_0^\infty = y$, and so $E[Y] = \int_0^\infty e^{-y}y\,dy = 1$. For the third, the inner integral is $\int_0^\infty xe^{-x/y}\,dx$ which we have already seen is y^2, and so $E[XY] = \int_0^\infty e^{-y}y^2 = e^{-y}(y^2+2y+2)\big|_0^\infty = 2$. Thus, $\mathrm{Cov}\,(X,Y) = E[XY] - E[X]E[Y] = 2 - 1 = 1$.

41. Because the variables are uncorrelated, $\mathrm{Cov}\,(X_i, X_j) = 0$ if $i \neq j$. Now, if $i \neq j$, $\mathrm{Var}(X_i + X_j) = \mathrm{Cov}\,(X_i + X_j, X_i + X_j) = \mathrm{Cov}\,(X_i, X_i) + 2\mathrm{Cov}\,(X_i, X_j) + \mathrm{Cov}\,(X_j, X_j) = 1 + 0 + 1 = 2$, and the standard deviation is $\sqrt{2}$.

(a) The correlation is $\mathrm{Cov}\,(X_1 + X_2, X_2 + X_3)$ divided by the product of the standard deviations, both of which are the $\sqrt{2}$. Now, $\mathrm{Cov}\,(X_1 + X_2, X_2 + X_3)$ can be expanded as the sum of four terms, three of which are 0, while the fourth is $\mathrm{Cov}\,(X_2, X_2) = 1$. Hence, the correlation is $\frac{1}{\sqrt{2}\sqrt{2}} = \frac{1}{2}$.

(b) When we expand $\mathrm{Cov}\,(X_1 + X_2, X_3 + X_4)$ we get the sum of four terms, all of which are 0, and so the correlation is 0 as well.

p. 421

1. This is $P(-20 \leq X - 20 \leq 20) = P(|X - 20| \leq 20)$. By Chebyshev's inequality, $P(|X - 20| > 20) \leq 20/20^2 = .05$, and so $P(|X - 20| \leq 20) \geq .95$.

2. (a) By Markov's inequality, $P(X > 85) \leq 75/85 = 15/17$.

(b) $P(65 \leq X \leq 85) = P(|X - 75| \leq 10) = 1 - P(|X - 75| > 10)$. By Chebyshev's inequality, $P(|X - 75| > 10) \leq 25/10^2 = \frac{1}{4}$, and so $P(65 \leq X \leq 85) \geq \frac{3}{4} = .75$.

(c) If we average n scores the mean is still 75 but the variance is n times the original variance divided by n^2, which has the effect of dividing the original variance by n. Thus, the new variance is 25/n, and by Chebyshev's inequality, if we call the average A_n, $P(|A_n - 75| \geq 5) \leq \frac{25/n}{5^2} = \frac{1}{n}$, and we want this to be at most .1, which is true if there are 10 or more students in the class.

3. We now work 2. (c) under the assumption that the distribution of the averaged scores is approximately normal. We have seen in 2. (c) that after averaging the mean is 75 and the variance is $\frac{25}{n}$, so that $\sigma = 5/\sqrt{n}$. Then $P(-70 \leq X \leq 80) = P(-5 \leq X - 75 \leq 5) = P(\frac{-5}{5/\sqrt{n}} \leq \frac{X-75}{5/\sqrt{n}} \leq \frac{5}{5/\sqrt{n}}) = P(-\sqrt{n} \leq \frac{X-75}{5/\sqrt{n}} \leq \sqrt{n})$ and if $\frac{X-75}{5/\sqrt{n}}$ is approximately unit normal this is $\Phi(\sqrt{n}) - \Phi(-\sqrt{n}) = \Phi(\sqrt{n}) - (1 - \Phi(\sqrt{n})) = 2\Phi(\sqrt{n}) - 1$, and this is at least .9 if $\Phi(\sqrt{n}) \geq (1 + .9)/2 = .95$. Now $\Phi(b) \geq .95$ for $b \geq 1.65$ from the table for Φ, and so we need $\sqrt{n} \geq 1.65$ of $n \geq (1.65)^2 = 2.7225$. This leads to the answer $n = 3$. However, it is not really clear that this value of n is large enough for the central limit theorem to give a valid approximation.