Department of Applied Mathematics and Statistics
The Johns Hopkins University

Introductory Examination–Summer Session

Wednesday, August 19, 2015

Instructions: Read carefully!

1. This **closed-book** examination consists of 15 problems, each worth 5 points. The passing grade has been set at 50 points, i.e., 2/3 of the total points. Partial credit will be given as appropriate; each part of a problem will be given the same weight. If you are unable to prove a result asserted in one part of a problem, you may still use that result to help in answering a later part.

2. You have been provided with a syllabus indicating the scope of the exam. Our purpose is to test not only your knowledge, but also your ability to apply that knowledge, and to provide mathematical arguments presented in **clear, logically justified steps**. The grading will reflect that broader purpose.

3. The problems have not been grouped by topic, but there are roughly equally many mainly motivated by each of the three areas identified in the syllabus (linear algebra; real analysis; probability). Nor have the problems been arranged systematically by difficulty. If a problem directs you to use a particular method of analysis, you *must* use it in order to receive substantial credit.

4. Start your answer to each problem on a NEW sheet of paper. Write only on ONE SIDE of each sheet, and please do not write very near the margins on any sheet. Arrange the sheets in order, and write your NAME and the PROBLEM NUMBER on each sheet.

5. The examination will begin at 8:30 AM; lunch and refreshments will be provided. The exam will end just before 5:00 PM. You may leave before then, but in that case you may not return.

6. Paper will be provided, but you should bring and use writing instruments that yield marks dark enough to be read easily.

7. **No calculators of any sort are needed or permitted.**

1. Consider families of n children, with $n \geq 2$. Let A be the event that a family has children of both sexes, and let B be the event that there is at most one girl in the family. Show that the only value of n for which the events A and B are independent is $n = 3$, assuming that each child has probability $1/2$ of being a boy.

Solution: We have

$$P(A) = 1 - P(A^c) = 1 - \big(P(\text{all boys}) + P(\text{all girls})\big) = 1 - (\frac{1}{2^n} + \frac{1}{2^n}) = \frac{2^{n-1} - 1}{2^{n-1}},$$

and

$$P(B) = P(\text{all boys}) + P(\text{one girl}) = \left(\frac{1}{2}\right)^n + n\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^{n-1} = \frac{n+1}{2^n},$$

while

$$P(A \cap B) = P(\text{one girl}) = \frac{n}{2^n}.$$

Since A and B are independent events if and only if $P(A \cap B) = P(A)P(B)$, we verify this condition if and only if

$$\frac{n}{2^n} = P(A)P(B) = \frac{2^{n-1} - 1}{2^{n-1}} \cdot \frac{n+1}{2^n} \iff 2^{n-1} = n + 1.$$

A direct substitution shows that this identity does not hold for $n = 2$ and holds for $n = 3$. We now prove by induction that $2^{n-1} > n + 1$ for $n \geq 4$. This is true for $n = 4$ since $8 = 2^{4-1} > 5 = 4 + 1$. Suppose $k \geq 4$ and $2^{k-1} > k + 1$. Then $2^k = 2(2^{k-1}) > 2(k + 1) > k + 2$, which completes the induction argument.

---

2. Suppose $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ are two power series having the same radius of convergence $\rho > 0$ and $\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} b_n x^n$ for $|x| < \rho$. Prove that the two series are identical, that is, $a_n = b_n$ for $n = 0, 1, 2, \ldots$.

Solution: Since $a_0 + a_1 x + a_2 x^2 + \cdots = b_0 + b_1 x + b_2 x^2 + \cdots$ when $|x| < \rho$, in particular, when $x = 0$ it immediately follows that $a_0 = b_0$. Furthermore, after subtracting this common term from each series we have $a_1 x + a_2 x^2 + a_3 x^3 + \cdots = b_1 x + b_2 x^2 + b_3 x^3 + \cdots$ for $|x| < \rho$. Clearly, each of these series has the same radius of convergence as the original series. We can write the last equivalence as

$$x(a_1 + a_2 x + a_3 x^2 + \cdots) = x(b_1 + b_2 x + b_3 x^2 + \cdots),$$

or, equivalently, if $x \neq 0$, as

$$a_1 + a_2x + a_3x^2 + \cdots = b_1 + b_2x + b_3x^2 + \cdots$$

when $0 < |x| < \rho$. It follows immediately that these series have radius of convergence ρ since

$$\limsup_{n\to\infty} |a_{n+1}|^{1/n} = \limsup_{n\to\infty} \left(|a_{n+1}|^{\frac{1}{n+1}}\right)^{\frac{n+1}{n}} = \frac{1}{\rho}.$$

Now, since $a_1 + a_2x + a_3x^2 + \cdots = b_1 + b_2x + b_3x^2 + \cdots$ for $0 < |x| < \rho$, by letting x tend to 0 through non-zero values, it follows that $a_1 = b_1$.

By continuing in this manner an inductive argument shows that $a_n = b_n$ for all $n = 0, 1, 2, \ldots$.

---

3. A four-digit number is selected at random. What is the probability that its leading digit is strictly larger than its second digit, its second digit is strictly larger than its third digit, and its third digit is strictly larger than its fourth digit. [Note that the leading digit of an n-digit number is nonzero.]

Solution: Use an equally-likely model, in which each of the possible 4-digit numbers are equally likely.

For the denominator, the leading digit can be $1, 2, 3, \ldots, 9$, but not 0, while each of the other digits can be $0, 1, 2, 3, \ldots, 9$, so by the basic counting principle, there are $9(10)^3 = 9000$ possible 4-digit numbers.

For the numerator, the 4 digits must be different, and must be in decreasing order. Thus, each four digit number in the event is obtained by choosing a subset of 4 or the 10 possible digits and placing them in decreasing order (the leading digit is then automatically nonzero).

The desired probability is then

$$\frac{\binom{10}{4}}{9000} = \frac{210}{9000} = \frac{7}{300} (= .023333\ldots).$$

---

4. Suppose that $f : [-1, 1] \to \mathbb{R}$ is a Riemann integrable function with

$$\int_{-1}^{1} f(x)^2 dx < \infty$$

and that

$$\int_{-1}^{1} f(x)\ dx = \int_{-1}^{1} f(x)(x^n + x^{n+1})\ dx$$

for all $n = 0, 1, 2, 3, ...$

Prove that $\int_{-1}^{1} f(x)g(x)dx = 0$ for all functions g continuous on $[-1, 1]$.

Solution: Let

$$\|f\|_2 = \sqrt{\int_{-1}^{1} f(x)^2 dx}.$$

From the Cauchy–Schwarz inequality, we have, for all $n \geq 0$,

$$\begin{aligned}
\left|\int_{-1}^{1} f(x)\ dx\right| &= \left|\int_{-1}^{1} f(x)(x^n + x^{n+1})\ dx\right| \\
&\leq \|f\|_2 \left(\int_{-1}^{1} (x^n + x^{n+1})^2 dx\right)^{1/2} \\
&= \|f\|_2 \left(\frac{2}{2n+1} + \frac{2}{2n+3}\right)^{1/2},
\end{aligned}$$

which implies that

$$\int_{-1}^{1} f(x)\ dx = 0.$$

The stated condition then implies that

$$\int_{-1}^{1} f(x)x^{n+1}\ dx = -\int_{-1}^{1} f(x)x^n\ dx$$

for all $n = 0, 1, 2, 3, ...$ This implies by induction that

$$\int_{-1}^{1} f(x)x^n\ dx = 0$$

for all $n = 0, 1, 2, 3, ...$ The Weierstrass approximation theorem states that for any continuous function g on $[-1, 1]$ and $\epsilon > 0$, there is a polynomial $p(x)$ such that $\max_{x\in[-1,1]} |g(x) - p(x)| < \epsilon$, so that, using Cauchy–Schwarz again,

$$\left|\int_{-1}^{1} f(x)g(x)\ dx\right| = \left|\int_{-1}^{1} f(x)[g(x) - p(x)]\ dx\right|$$
$$\leq \|f\|_2 \left(\int_{-1}^{1} [g(x) - p(x)]^2 dx\right)^{1/2} \leq \sqrt{2}\epsilon\|f\|_2.$$

Because ϵ is arbitrary,

$$\int_{-1}^{1} f(x)g(x)\ dx = 0$$

for any continuous function g.

---

5. Let $A \in \mathbb{R}^{n\times n}$ be a symmetric positive semidefinite matrix for some $n \geq 1$. Prove that $x^T Ax = 0$ if and only if $Ax = 0$.

Solution: If $Ax = 0$, then clearly $x^T Ax = 0$, which proves the "if" direction.

For the "only if" direction, suppose that $x^T Ax = 0$. Since A is symmetric and positive semidefinite, we know that x can be represented in a basis consisting of orthonormal eigenvectors, i.e., that

$$x = \sum_{i=1}^{n} \alpha_i v_i = \sum_{i\in\Lambda^+} \alpha_i v_i + \sum_{i\in\Lambda^0} \alpha_i v_i \tag{1}$$

for some scalars $\{\alpha_i\}_{i=1}^n$, where $\{v_i\}_{i=1}^n$ are a set of orthonormal eigenvectors for A with associated eigenvalues $\{\lambda_i\}_{i=1}^n$, $\Lambda^+ = \{i : \lambda_i > 0\}$, and $\Lambda^0 = \{i : \lambda_i = 0\}$. (Recall that all eigenvalues of A are nonnegative since A is positive semidefinite by assumption.) Using (**??**) and the orthogonality of the eigenvectors $\{v_i\}$, we may conclude that

$$0 = x^T Ax = \sum_{i=1}^{n} \alpha_i^2 \lambda_i = \sum_{i\in\Lambda^+} \alpha_i^2 \lambda_i + \sum_{i\in\Lambda^0} \alpha_i^2 \lambda_i = \sum_{i\in\Lambda^+} \alpha_i^2 \lambda_i. \tag{2}$$

Since $\lambda_i > 0$ for $i \in \Lambda^+$, we know from (**??**) that $\alpha_i = 0$ for all $i \in \Lambda^+$, which when combined with the identity in (**??**) means that

$$x = \sum_{i\in\Lambda^0} \alpha_i v_i. \tag{3}$$

Since $Av_i = \lambda_i v_i = 0$ for all $i \in \Lambda^0$ (i.e., $v_i \in \text{Null}(A)$ for all $i \in \Lambda^0$), we may conclude from (**??**) that $x \in \text{Null}(A)$, i.e., that $Ax = 0$. This completes the proof.

---

6. Let $X \in \mathbb{R}^{D\times N}$ for some positive integers D and N, and $Z \in \mathbb{R}^{D\times N}$ be a random matrix whose ND entries are independent, each z_{ij} being a Bernoulli random variable with parameter p, i.e.,

$$P[z_{ij} = k] = (1-p)^{1-k} p^k \quad \text{for } k \in \{0, 1\},$$

and for all $i = 1, \ldots, D$ and $j = 1, \ldots, N$. Prove that the matrix

$$\Gamma := Y^T Y + (p-1)\mathrm{diag}(Y^T Y) \;\text{ with }\; Y := \frac{1}{p}(Z \odot X)$$

satisfies $\mathbb{E}[\Gamma] = X^T X$, where $\odot$ denotes the entry-wise product of matrices and

$$(\mathrm{diag}(Y^T Y))_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ (Y^T Y)_{ij} & \text{if } i = j. \end{cases}$$

Solution: Let y_i denote the ith column of Y. We consider two cases.
If $j \neq i$, then

$$\Gamma_{ij} = (Y^T Y)_{ij} = y_i^T y_j = \frac{1}{p^2} \sum_{k=1}^{D} x_{ki} x_{kj} z_{ki} z_{kj}.$$

It follows that

$$\mathbb{E}[\Gamma_{ij}] = \frac{1}{p^2} \sum_{k=1}^{D} x_{ki} x_{kj} \mathbb{E}[z_{ki} z_{kj}] = \sum_{k=1}^{D} x_{ki} x_{kj} = x_i^T x_j = (X^T X)_{ij},$$

where we used $E[z_{ki} z_{kj}] = \mathbb{E}[z_{ki}]\mathbb{E}[z_{kj}]$ because $z_{k,i}$ and $z_{k,j}$ are independent, and $\mathbb{E}[z_{ki}] = 0 \cdot (1-p) + 1 \cdot p = p$.
If $j = i$, then

$$\Gamma_{ii} = (Y^T Y)_{ii} + (p-1)(Y^T Y)_{ii} = p \cdot (Y^T Y)_{ii} = p \sum_{k=1}^{D} y_{ki} y_{ki} = \frac{1}{p} \sum_{k=1}^{D} x_{ki}^2 z_{ki}^2.$$

It follows that

$$\mathbb{E}[\Gamma_{ii}] = \frac{1}{p} \sum_{k=1}^{D} x_{ki}^2 \mathbb{E}[z_{ki}^2] = (X^T X)_{ii},$$

where we use the fact that $\mathbb{E}[z_{ki}^2] = 0 \cdot (1-p) + 1 \cdot p = p$.
The desired result follows from the above two cases.

---

7. Let A be a real $n \times n$ positive definite matrix. Prove that $\det(A) \leq (\text{trace}(A)/n)^n$.

   Hint: Use the arithmetic-geometric mean inequality.

   Solution: Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the n positive eigenvalues of A. By the arithmetic-geometric mean inequality, $(\prod_{i=1}^n \lambda_i)^{1/n} \leq \sum_{i=1}^n \lambda_i/n$. The result then follows by noting that $\det(A) = \prod_{i=1}^n \lambda_i$ and $\text{trace}(A) = \sum_{i=1}^n \lambda_i$.

---

8. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by setting $f(0,0) := 0$ and

$$f(x,y) := \frac{xy(x^2 - y^2)}{x^2 + y^2} \text{ if } (x,y) \neq (0,0).$$

   Prove that $\partial_1 \partial_2 f(0,0)$ and $\partial_2 \partial_1 f(0,0)$ exist, and that they are not equal, where ∂_1 is the partial derivative with respect to the first variable (x) and ∂_2 with respect to the second one (y).

   Solution:

   Since f is a ratio of non-vanishing polynomials for $(x,y) \neq (0,0)$, it is infinitely differentiable everywhere except possible at $(0,0)$. We first prove the existence of first derivatives at $(0,0)$:

$$\partial_1 f(0,0) = \lim_{x \to 0} \frac{f(x,0) - f(0,0)}{x - 0} = \lim_{x \to 0} \frac{0 - 0}{x - 0} = 0$$

   and similarly $\partial_2 f(0,0)$ exists and is equal to 0. Similarly,

$$\partial_1 f(0,y) = \lim_{x \to 0} \frac{f(x,y) - f(0,y)}{x - 0} = \frac{-y^3}{y^2} = -y.$$

   Therefore $\partial_2 \partial_1 f(0,0)$ exists and equals -1. Since $f(x,y) \equiv -f(y,x)$, we have

$$\partial_2 f(x,0) = -\partial_1 f(0,x) = x \text{ for } x \neq 0 \text{ so that}$$

$$\partial_1 \partial_2 f(0,0) = 1.$$

---

9. (All matrices in this problem are assumed to be real.) Consider symmetric $n \times n$ matrices A and B, and assume that $B - A$ is positive definite. Find necessary and sufficient conditions on $n \times n$ matrices C such that $CBC^T - CAC^T$ is also positive definite.

Solution: It is necessary and sufficient that C is an invertible matrix. Indeed, since $B-A$ is symmetric positive definite, there exists an $n \times n$ invertible matrix E such that $B - A = EE^T$. Thus $CBC^T - CAC^T = C(B-A)C^T = CEE^TC^T = (CE)(CE)^T$. This is always positive semidefinite, and positive definite if and only if CE, hence C, is invertible.

---

10. Let $f : \mathbb{R}^+ \to \mathbb{R}$ be a function that satisfies the following conditions:

    (a) f is continuous,

    (b) f satisfies $f(xy) = f(x) + f(y)$ for all $x, y > 0$,

    (c) $f(1) = 0$, and

    (d) $f(e) = 1$.

    Prove that $f(x) = \ln x$.

    Hint: Begin by considering x values that are integer powers of e, i.e., $x = e^a$ for $a \in \mathbb{Z}$ and then rational powers of e, i.e., $x = e^{a/b}$ for $a, b \in \mathbb{Z}$, $b \neq 0$.

Solution: For a positive integer a, we can apply (b) repeatedly to deduce

$$f(e^a) = f(e \cdot e \cdot e \cdots e) = f(e) + f(e) + \cdots + f(e) = 1 + 1 + \cdots + 1 = a.$$

From this, we have

$$0 = f(1) = f(e^a \cdot e^{-a}) = f(e^a) + f(e^{-a}) = a + f(e^{-a})$$

and so $f(e^{-a}) = -a$.

Therefore, for any integer a (positive, negative, or zero) we have $f(e^a) = a$.

Now let b be a positive integer and a be any integer. We have

$$a = f(e^a) = f\left(\underbrace{e^{a/b} \cdot e^{a/b} \cdots e^{a/b}}_{b \text{ times}}\right) = f\left(e^{a/b}\right) + f\left(e^{a/b}\right) + \cdots + f\left(e^{a/b}\right) = bf\left(e^{a/b}\right)$$

which implies $f\left(e^{a/b}\right) = a/b$.

Thus, for any $r \in \mathbb{Q}$, $f(e^r) = r$.

Finally, let x be any positive real number and choose a sequence of rational numbers $r_1, r_2, r_3, \ldots$ that converges to $\ln x$. By the continuity of f we have

$$f(x) = f\left(e^{\ln x}\right) = \lim_{n\to\infty} f\left(e^{r_n}\right) = \lim_{n\to\infty} r_n = \ln x.$$

---

11. Let $\alpha = 1 + \sqrt{2}$. Because $\alpha > 1$, we know that α^n diverges as $n \to \infty$. However, if we look at the values produced, it is interesting to note that α^n gets closer and closer to being an integer. For example,

$$\left(1 + \sqrt{2}\right)^{20} = 45239073.999999977895215\ldots$$

Explain why this is so, that is, prove that there exists a sequence z_n of integers such that

$$\lim_{n\to\infty} \left[\alpha^n - z_n\right] = 0.$$

Hint: Find β such that $\alpha^n + \beta^n$ is an integer for all n.

Solution: Let $\beta = 1 - \sqrt{2} \approx -0.414$, so $\beta^n \to 0$ as $n \to \infty$. Thus the difference between α^n and $\alpha^n + \beta^n$ goes to zero. By the Binomial Theorem (twice), we have

$$z_n := \alpha^n + \beta^n = \sum_{k=0}^{n} \binom{n}{k} \left(\sqrt{2}^k + (-1)^k \sqrt{2}^k\right). \qquad (*)$$

In $(*)$, when k is odd, the summands are zero. When k is even, the summands are integers. Therefore z_n, which grows arbitrarily close to α^n, is an integer for all n.

---

12. Let $\ldots, Z_{-2}, Z_{-1}, Z_0, Z_1, Z_2, \ldots$ be a (doubly) infinite sequence of independent identically distributed standard normal random variables. For each integer n, let $Y_n = Z_n Z_{n-1}$. Show that the sequence (Y_n) is uncorrelated and *not* pairwise independent. That is, show for all integers $i \neq j$, $E(Y_iY_j) = E(Y_i)E(Y_j)$, but that there is some integer $i \neq j$ such that Y_i and Y_j are not independent.

Solution: In the solution below we will use the facts that for standard normal random variables Z we have $E(Z) = 0$, $E(Z^2) = 1$, and $E(Z^4) = 3$.

First of all, since $E(Y_n) = E(Z_n Z_{n-1}) = E(Z_n)E(Z_{n-1}) = 0$ for all n, to show (Y_n) is uncorrelated it is enough to show for integers $i \neq j$ that $E(Y_i Y_j) = 0$. To this end we can assume without loss of generality that $i > j$. Therefore,

$$\begin{aligned} E(Y_i Y_j) &= E(Z_i Z_{i-1} Z_j Z_{j-1}) \\ &= E(Z_i) E(Z_{i-1} Z_j Z_{j-1}) = 0, \end{aligned}$$

and (Y_n) is uncorrelated.

To see that (Y_n) is not a pairwise independent sequence, we will show that $E(Y_2^2 Y_1^2) \neq E(Y_2^2)E(Y_1^2)$. To this end note that for any integer i, $E(Y_i^2) = E(Z_i^2 Z_{i-1}^2) = E(Z_i^2)E(Z_{i-1}^2) = 1$. Moreover,

$$\begin{aligned} E(Y_2^2 Y_1^2) &= E(Z_2^2 Z_1^4 Z_0^2) \\ &= E(Z_2^2)E(Z_1^4)E(Z_0^2) = 3. \end{aligned}$$

Therefore, since $E(Y_2^2)E(Y_1^2) = 1 \neq 3 = E(Y_2^2 Y_1^2)$, the sequence (Y_n) cannot be a pairwise independent sequence.

---

13. Let F and G be two subspaces of $\mathbb{R}^n$ such that $\dim(F) + \dim(G) = n$. Prove that there exists a linear map $A : \mathbb{R}^n \to \mathbb{R}^n$ such that $\mathrm{Null}(A) = F$ and $\mathrm{Range}(A) = G$.

Solution: Let $m = \dim(F)$ and $k = \dim(G)$. Let $(e_1, \ldots, e_m, e_{m+1}, \ldots, e_{m+k})$ be a basis of $\mathbb{R}^n$ such that $(e_1, \ldots, e_m)$ is a basis for F. Similarly, let $(f_1, \ldots, f_k)$ be a basis of G. Define A by $Ae_i = 0$ for $i \in \{1, \ldots, m\}$ and $Ae_{m+j} = f_j$ for $j \in \{1, \ldots, k\}$. In other terms, A is the linear transformation with matrix

$$[\underbrace{0, \ldots, 0}_{m \text{ times}}, f_1, \ldots, f_k] \; [e_1, \ldots, e_n]^{-1}$$

relative to the basis $(e_1, \ldots, e_n)$.

Then $\mathrm{Null}(A) \supset F$ and $\mathrm{Range}(A) \supset G$. But if one of these inclusions were strict, we would have

$$n = \dim(\mathrm{Range}(A)) + \dim(\mathrm{Null}(A)) > k + m = n,$$

which is a contradiction.

---

14. Let S be an n-by-n (real) matrix with rank m. Prove that there exist two real matrices A and B such that A is n-by-m, B is m-by-n, and $S = AB$.

Solution: Note that one must have $m \leq n$. Let $e_1, \ldots, e_m \in \mathbb{R}^n$ be a basis of $\text{Range}(S)$. Take $A = [e_1, \ldots, e_m]$. Since A has rank m, $A^T A$ is invertible, and if B is such that $S = AB$, we must have $A^T S = A^T AB$, leaving

$$B = (A^T A)^{-1} A^T S$$

as the only possible choice. Taking this B, one then has $AB = S$. Indeed, if $x \in \mathbb{R}^n$, then $Sx \in \text{Range}(S)$ is a linear combination of the columns of A, i.e., there exists $\lambda \in \mathbb{R}^m$ such that $Sx = A\lambda$. One then has

$$ABx = A(A^T A)^{-1} A^T Sx = A(A^T A)^{-1} A^T A\lambda = A\lambda = Sx.$$

---

15. A football team consists of 20 offensive and 20 defensive players. The players are to be grouped in groups of 2 for the purpose of determining roommates. If the pairing is done (uniformly) at random, what is the probability that there are no offensive–defensive roommate pairs? Express your answer as a ratio of products of factorials.

Solution: #1: This is an Example in Chapter 2 of Ross. There are

$$\binom{40}{2, 2, \ldots, 2} = \frac{40!}{(2!)^{20}}$$

ways of dividing the 40 players into 20 *ordered* pairs of two each. Hence there are $40!/[(2!)^{20} 20!]$ ways of dividing the players into (unordered) pairs of 2 each. Furthermore, since a division will result in no offensive–defensive pairs if and only if the offensive (and defensive) players are paired among themselves, it follows that there are $\{20!/[(2!)^{10} 10!]\}^2$ such divisions. Hence the probability of no offensive–defensive roommate pairs is given by

$$\frac{\{20!/[(2!)^{10} 10!]\}^2}{40!/[(2!)^{20} 20!]} = \frac{(20!)^3}{(10!)^2 40!} \left[= \frac{323}{240990435} \approx 1.34 \times 10^{-6} \right].$$

#2: Label the offensive players as $O_1, \ldots, O_{20}$. Consider a sequence of steps at which the lowest-numbered not-yet-paired offensive player is chosen and paired with a randomly chosen not-yet-paired player. For $i = 1, \ldots, 10$, the conditional probability

that, at the ith such step, the offensive player chosen is paired with an offensive player given that the same is true at steps $1, 2, \ldots, i-1$ is clearly $[20-(2i-1)]/[40-(2i-1)] = (21-2i)/(41-2i)$. Thus the probability that there are no offensive–defensive pairs is

$$\prod_{i=1}^{10} \frac{21-2i}{41-2i} = \frac{[20!/(2^{10}10!)]^2}{40!/(2^{20}20!)} = \frac{(20!)^3}{(10!)^2 40!}.$$
