# EE5137 Stochastic Processes: Problem Set 2
## Assigned: 22/01/21, Due: 29/01/21

There are five non-optional problems in this problem set.

1. Exercise 1.12 (Gallager's book) Let $X$ be a rv with CDF $F_X(x)$. Find the CDF of the following rv's.

   (a) The maximum of $n$ IID rv's, each with CDF $F_X(x)$.

   (b) The minimum of $n$ IID rv's, each with CDF $F_X(x)$.

   (c) The difference of the rv's defined in a) and b); assume $X$ has a density $f_X(x)$.

   *Hint: For parts (a) express the CDF of $M_+$ (the maximum of the $N$ rvs) in terms of the CDFs of the individual rvs. Part (b) is analogous. Part (c) is most challenging. You may first condition on the event $\{X_1 = x\}$. Then note that $X_1 = M_+$ iff $X_j \le x$ for all $2 \le j \le n$. Also given $X_1 = M_+ = x$, we have $R = M_+ - M_- \le r$ iff $X_j > x - r$ for $2 \le j \le n$. Now since the rvs are i.i.d.,*

   $$\Pr(M_+ = X_1, R \le r \mid X_1 = x) = \prod_{j=2}^{n} \Pr(x - r < X_j \le x)$$

   *Continue the above argument (average over $X_1 = x$) to show that*

   $$\Pr(R \le r) = \int_{-\infty}^{\infty} n f_X(x)[F_X(x) - F_X(x-r)]^{n-1}\,\mathrm{d}x.$$

   **Solution:**

   (a) Let $M_+$ be the maximum of the $n$ rv's $X_1, X_2, \ldots, X_n$. Note that for any real $x$, $M_+$ is less than or equal to $x$ if and only if $X_j \le x$ for each $j, 1 \le j \le n$. Thus

   $$\Pr\{M_+ \le x\} = \Pr\{X_1 \le x, X_2 \le x, \cdots, X_n \le x\} = \prod_{j=1}^{n} \Pr\{X_j \le x\}, \tag{1}$$

   where we have used the independence of the $X_j$'s. Finally, since $\Pr\{X_j \le x\} = F_X(x)$ for each $j$, we have $F_{M_+}(x) = \Pr\{M_+ \le x\} = (F_X(x))^n$.

   (b) Let $M_-$ be the minimum of $X_1, X_2, \ldots, X_n$. Then, in the same way as in $((a), M_- > y$ if and only if $X_j > y$ for $1 \le j \le n$ and for all choice of $y$. We could make the same statement using greater than or equal in place of strictly greater than, but the strict inequality is what is needed for the CDF. Thus,

   $$\Pr\{M_- > y\} = \Pr\{X_1 > y, X_2 > y, \ldots, X_n > y\} = \prod_{j=1}^{n} \Pr\{X_j > y\}. \tag{2}$$

   It follows that $1 - F_{M_-}(y) = (1 - F_X(y))^n$.

(c) There are many difficult ways to do this, but also a simple way, based on first conditioning on the event that $X_1 = x$. Then $X_1 = M_+$ if and only if $X_j \le x$ for $2 \le j \le n$. Also, given $X_1 = M_+ = x$, we have $R = M_+ - M_- \le r$ if and only if $X_j > x - r$ for $2 \le j \le n$. Thus, since the $X_j$ are IID,

$$\Pr\{M_+ = X_1, R \le r | X_1 = x\} = \prod_{j=2}^{n} \Pr\{x - r < X_j \le x\} \tag{3}$$

$$= [\Pr\{x - r < X \le x\}]^{n-1} \tag{4}$$

$$= [F_X(x) - F_X(x - r)]^{n-1}. \tag{5}$$

We can now remove the conditioning by averaging over $X_1 = x$. Assuming that $X$ has the density $f_X(x)$,

$$\Pr\{X_1 = M_+, R \le r\} = \int_{-\infty}^{\infty} f_X(x)[F_X(x) - F_X(x - r)]^{n-1} dx. \tag{6}$$

Finally, we note that the probability that two of the $X_j$ are the same is 0 so the events $X_j = M_+$ are disjoint except with zero probability. Also we could condition on $X_j = x$ instead of $X_1$ with the same argument (i.e., by using symmetry), so $\Pr\{X_j = M_+, R \le r\} = \Pr\{X_1 = M_+, R = r\}$. It follows that

$$\Pr\{R \le r\} = \int_{-\infty}^{\infty} n f_X(x)[F_X(x) - F_X(x - r)]^{n-1} dx. \tag{7}$$

2. Exercise 1.14 (Gallager's book)

   (a) Let $X_1, X_2, \dots, X_n$ be rv's with expected values $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n$. Show that $\mathbb{E}[X_1 + X_2 + \cdots + X_n] = \bar{X}_1 + \bar{X}_2 + \cdots + \bar{X}_n$. You may assume that the rv's has a joint density function, but do not assume that the rv's are independent.

   (b) Now assume that $X_1, X_2, \cdots, X_n$ are statically independent and show that the expected value of the product is equal to the product of the expected values.

   (c) Again assuming that $X_1, X_2, \cdots, X_n$ are statistically independent, show that the variance of the sum is equal to the sum of the variances.

**Solution:**

(a) We assume that the rv's have a joint density, and we ignore all mathematical fine points here. Then

$$\mathbb{E}[X_1 + X_2 + \cdots + X_n] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1 + x_2 + \cdots + x_n) f_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n \tag{8}$$

$$= \sum_{j=1}^{n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_j f_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n \tag{9}$$

$$= \sum_{j=1}^{n} \int_{-\infty}^{\infty} x_j f_{X_j}(x_j) dx_j \tag{10}$$

$$= \sum_{j=1}^{n} \mathbb{E}[X_j]. \tag{11}$$

Note that the separation into a sum of integrals simply used the properties of integration and that no assumption of statistical independence was made.

(b) From the independence, $f_{X_1X_2\cdots X_n}(x_1, x_2, \cdots, x_n) = \prod_{j=1}^n f_{X_j}(x_j)$. Thus,

$$\mathbb{E}[X_1X_2\cdots X_n] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{j=1}^{n} x_j \prod_{j=1}^{n} f_{X_j}(x_j)dx_1\cdots dx_n \tag{12}$$

$$= \prod_{j=1}^{n} \int_{-\infty}^{\infty} x_j f_{X_j}(x_j)dx_j \tag{13}$$

$$= \prod_{j=1}^{n} \mathbb{E}[X_j]. \tag{14}$$

(c) Since (a) shows that $\mathbb{E}[\sum_{j=1}^n X_j] = \sum_{j=1}^n \bar{X}_j$, we have

$$\mathsf{Var}\left[\sum_{j=1}^{n} X_j\right] = \mathbb{E}\left[\left(\sum_{j=1}^{n} X_j - \sum_{j=1}^{n} \bar{X}_j\right)^2\right] \tag{15}$$

$$= \mathbb{E}\left[\sum_{j=1}^{n}\sum_{i=1}^{n}(X_j - \bar{X}_j)(X_i - \bar{X}_i)\right] \tag{16}$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{n}\mathbb{E}\left[(X_j - \bar{X}_j)(X_i - \bar{X}_i)\right], \tag{17}$$

where we have again used $(a)$. Now from $(b)$ (which used the independent of the $X_j$), $\mathbb{E}[(X_j - \bar{X}_j)(X_i - \bar{X}_i)] = 0$ for $i \neq j$. Thus (17) simplifies to

$$\mathsf{Var}\left[\sum_{j=1}^{n} X_j\right] = \sum_{j=1}^{n}\mathbb{E}[(X_j - \bar{X}_j)^2] = \sum_{j=1}^{n}\mathsf{Var}[X_j]. \tag{18}$$

3. Exercise 1.20 (Gallager's book)

(a) Consider a positive, integer-valued rv whose CDF is given at integer values by

$$F_Y(y) = 1 - \frac{2}{(y+1)(y+2)}, \tag{19}$$

for integer $y \geq 0$. Use (1.31) to show that $\mathbb{E}[Y] = 2$. Hint: note that $1/[(y+1)(y+2)] = 1/(y+1) - 1/(y+2)$.

(b) Find the PMF of $Y$ and use it to check the value of $\mathbb{E}[Y]$.

(c) Let $X$ be another positive, integer-valued rv. Assume its conditional PDF is given by

$$p_{X|Y}(x|y) = \frac{1}{y}, \quad \text{for} \quad 1 \leq x \leq y. \tag{20}$$

Find $\mathbb{E}[X|Y = y]$ and use it to show that $\mathbb{E}[X] = 3/2$. Explore finding $p_X(x)$ until you are convinced that using the conditional expectation to calculate $\mathbb{E}[X]$ is considerably esier than using $p_X(x)$.

(d) Let $Z$ be another integer-valued rv with the conditional PMF

$$p_{Z|Y}(z|y) = \frac{1}{y^2}, \quad \text{for} \quad 1 \leq z \leq y^2. \tag{21}$$

Find $\mathbb{E}[Z|Y = y]$ for each integer $y \geq 1$ and find $\mathbb{E}[Z]$.

**Solution:**

(a) Combining (1.31) with the hint, we have

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{y\ge 0} F_Y^c(y) && (22)\\
&= \sum_{y\ge 0}\frac{2}{y+1} - \sum_{y\ge 0}\frac{2}{y+2} && (23)\\
&= \sum_{y\ge 0}\frac{2}{y+1} - \frac{2}{y+2} && (24)\\
&= 2, && (25)
\end{aligned}$$

where the second sum in the second line eliminates all but the first term of the first sum.

(b) For $y = 0, p_Y(y) = F_Y(y) = 0$. For integer $y \ge 1$, $p_Y(y) = F_Y(y) - F_Y(y-1)$. Thus for $y \ge 1$,

$$p_Y(y) = \frac{2}{y(y+1)} - \frac{2}{(y+1)(y+2)} = \frac{4}{y(y+1)(y+2)}. \tag{26}$$

Find $\mathbb{E}[Y]$ from the PMF, we have

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{y=1}^{\infty} y p_Y(y) && (27)\\
&= \sum_{y=1}^{\infty}\frac{4}{(y+1)(y+2)} && (28)\\
&= \sum_{y=1}^{\infty}\frac{4}{y+1} - \sum_{y=2}^{\infty}\frac{4}{y+1} && (29)\\
&= 2. && (30)
\end{aligned}$$

(c) Conditioned on $Y = y$, $X$ is uniform over $\{1, 2, \cdots, y\}$ and thus has the conditional mean $(y+1)/2$. It follows that

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}\left[\frac{Y+1}{2}\right] = \frac{3}{2}. \tag{31}$$

Calculating this expectation in the conventional way would require first calculating $p_X(x)$ and then calculating the expectation. Calculating $p_X(x)$,

$$p_X(x) = \sum_{y=x}^{\infty} p_Y(y) p_{X|Y}(x|y) = \sum_{y=x}^{\infty}\frac{4}{y(y+1)(y+2)} \times \frac{1}{y}. \tag{32}$$

(d) As in (c), $\mathbb{E}[Z|Y] = (Y^2+1)/2$. Since $p_Y(y)$ approaches 0 as $y^{-3}$, we see that $\mathbb{E}[Y^2]$ is infinite and thus $\mathbb{E}[Z] = \infty$.

4. Exercise 1.22 (Gallager's book)

Suppose $X$ has the Poisson PMF, $p_X(n) = \lambda^n \exp(-\lambda)/n!$ for $n \ge 0$ and $Y$ has the Poisson PMF, $p_Y(n) = \mu^n \exp(-\mu)/n!$ for $n \ge 0$. Assume that $X$ and $Y$ are independent. Find the distribution of $Z = X + Y$ and find the conditional distribution of $Y$ conditional on $Z = n$.

**Solution:** The seemingly straightforward approach is to take the discrete convolution of $X$ and $Y$ (i.e., the sum of the joint PMF's of $X$ and $Y$ for which $X+Y$ has a given value $Z=n$). Thus

$$p_Z(n) = \sum_{k=0}^{n} p_X(k)p_Y(n-k) = \sum_{k=0}^{n} \frac{\lambda^k e^{-\lambda}}{k!} \cdot \frac{\mu^{n-k}e^{-\mu}}{(n-k)!} \tag{33}$$

$$= e^{-(\lambda+\mu)} \sum_{k=0}^{n} \frac{\lambda^k \mu^{n-k}}{k!(n-k)!}. \tag{34}$$

At this point, one needs some added knowledge or luck. One might hypothesize (correctly) that $Z$ is also a Poisson rv with parameter $\lambda+\mu$; one might recognize the sum above, or one might look at an old solution. We multiply and divide the right hand expression above by $(\lambda+\mu)^n/n!$.

$$p_Z(n) = \frac{(\lambda+\mu)^n e^{-(\lambda+\mu)}}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(\frac{\mu}{\lambda+\mu}\right)^{n-k} \tag{35}$$

$$= \frac{(\lambda+\mu)^n e^{-(\lambda+\mu)}}{n!}. \tag{36}$$

where we have recognized the sum on the right as a binomial sum.

Another approach that is actually more straightforward uses the fact (see (1.52)) that the MGF of the sum of independent rv's is the product of the MGF's of those rv's. From Table 1.2 (or a simple derivation), $g_X(r) = \exp[\lambda(e^r-1)]$. Similarly, $g_Y(r) = \exp[\mu(e^r-1)]$. Hence, $g_Z(r) = \exp[(\lambda+\mu)(e^r-1)]$. Since the MGF specifies the PMF, $Z$ is a Poisson rv with parameter $\lambda+\mu$.

Finally, we must find $p_{Y|Z}(i|n)$. As a prelude to using Bayes' law, note that

$$p_{Z|Y}(n|i) = \Pr(X+Y=n|Y=i) = \Pr(X=n-i). \tag{37}$$

Thus,

$$p_{Y|Z}(i|n) = \frac{p_Y(i)p_X(n-i)}{p_Z(n)} = \frac{\lambda^{n-i}e^{-\lambda}}{(n-i)!} \cdot \frac{\mu^i e^{-\mu}}{i!} \cdot \frac{n!}{(\mu+\lambda)^n} e^{-(\lambda+\mu)} \tag{38}$$

$$= \binom{n}{i} \left(\frac{\lambda}{\lambda+\mu}\right)^{n-i} \left(\frac{\mu}{\lambda+\mu}\right)^{i}. \tag{39}$$

Why this turns out to be a binomial PMF will be clarified when we study Poisson processes.

5. (a) Using the law of iterated expectations,

$$\mathbb{E}[X_i] = \mathbb{E}[\mathbb{E}[X_i|Y]] = \mathbb{E}[Y] = \mu$$

where the second equality is because $\mathbb{E}[X_i|Y=y]=y$. Now, the expectation of the sum is

$$\mathbb{E}[S_n] = \mathbb{E}[X_1 + \ldots + X_n] = n\mu.$$

(b) The variance is

$$\mathrm{Var}(X_i) = \mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2 = \mathbb{E}[X_i] - (\mathbb{E}[X_i])^2 = \mu - \mu^2.$$

(c) The covariance can be computed as

$$\begin{aligned}
\mathrm{Cov}(X_i, X_j) &= \mathbb{E}[X_iX_j] - \mathbb{E}[X_i]\mathbb{E}[X_j] \\
&= \mathbb{E}[\mathbb{E}[X_iX_j|Y]] - \mu^2 \\
&= \mathbb{E}[\mathbb{E}[X_i|Y]\mathbb{E}[X_j|Y]] - \mu^2 \\
&= \mathbb{E}[Y^2] - \mu^2 \\
&= \sigma^2 + \mu^2 - \mu^2 \\
&= \sigma^2 > 0.
\end{aligned}$$

Hence, the random variables $X_i$ and $X_j$ are not independent.

(d) We now derive $\mathrm{Var}(S_n) = \mathbb{E}[S_n^2] - (\mathbb{E}[S_n])^2$ using the law of iterated expectations. We have

$$\begin{aligned}\mathbb{E}[S_n^2] - \mathbb{E}[S_n]^2 &= \mathbb{E}[\mathbb{E}[S_n^2|Y]] - (\mathbb{E}[\mathbb{E}[S_n|Y]])^2 \\ &= \mathbb{E}[\mathbb{E}[S_n^2|Y] - \mathbb{E}[S_n|Y]^2] + \mathbb{E}[\mathbb{E}[S_n|Y]^2] - (\mathbb{E}[\mathbb{E}[S_n|Y]])^2 \\ &= \mathbb{E}[\mathrm{Var}(S_n|Y)] + \mathrm{Var}(\mathbb{E}[S_n|Y])\end{aligned}$$

(e) We calculate $\mathrm{Var}(S_n|Y)$ first. By conditional independence, we have

$$\begin{aligned}\mathrm{Var}(S_n|Y) &= \mathrm{Var}(X_1 + \ldots + X_n|Y) \\ &= nY(1-Y)\end{aligned}$$

Next we recall that $\mathbb{E}[S_n|Y] = nY$. Thus,

$$\begin{aligned}\mathrm{Var}(S_n) &= \mathbb{E}[nY(1-Y)] + \mathrm{Var}(nY) \\ &= n[\mu - \sigma^2 - \mu^2] + n^2\sigma^2 \\ &= n[\mu - \mu^2] + n(n-1)\sigma^2.\end{aligned}$$

---

6. (Optional) Exercise 1.6 (Gallager's book) We have $\Pr(X > x) = \int_x^\infty f_X(y)\,\mathrm{d}y$ from the definition of a continuous rv. We look at $\mathbb{E}[X] = \int_0^\infty \Pr(X > x)\,\mathrm{d}x$ as $\lim_{a\to\infty} \int_0^a F_X^c(x)\,\mathrm{d}x$ since the limiting operation $a \to \infty$ is where the interesting issue is.

$$\begin{aligned}\int_0^a F_X^c(x)\,\mathrm{d}x &= \int_0^a \int_x^\infty f_X(y)\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_0^a \int_x^a f_X(y)\,\mathrm{d}y\,\mathrm{d}x + \int_0^a \int_a^\infty f_X(y)\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_0^a \int_0^y f_X(y)\,\mathrm{d}x\,\mathrm{d}y + aF_X^c(a).\end{aligned}$$

We first broke the integral on the right into two parts, one for $y < x$ and the other for $y \geq x$. Since the limits of integration on the first part were finite, they could be interchanged. The inner integral of the first part is $yf_X(y)$, so

$$\lim_{a\to\infty} \int_0^a F_X^c(x)\,\mathrm{d}x = \lim_{a\to\infty} \int_0^a yf_X(y)\,\mathrm{d}y + \lim_{a\to\infty} aF_X^c(a).$$

Assuming that $\mathbb{E}[X]$ exists, the integral on the left is nondecreasing in $x$ and has the finite limit $\mathbb{E}[X]$. The first integral on the right is also nondecreasing and upper bounded by the first integral, so it also has a limit. This means that $\lim_{a\to\infty} aF_X^c(a)$ must also have a limit, say $\beta$. Now if $\beta > 0$, then for any $\epsilon \in (0,\beta)$, $aF_X(a) > \beta - \epsilon$ for all sufficiently large $a$. For all such $a$, then $F_X^c(a) > (\beta-\epsilon)/a$. This would imply that $\mathbb{E}[X] = \int_0^\infty F_X^c(x)\,\mathrm{d}x = \infty$, which is a contradiction. Thus $\beta = 0$, i.e., $\lim_{a\to\infty} aF_X^c(a) = 0$, establishing the claim.

7. (Optional) Exercise 1.16 (Gallager's book) Let $X_1, X_2, \cdots, X_n$ be a sequence of IID continuous rv's with the common probability density function $f_X(x)$; note that $\Pr\{X = \alpha\} = 0$ for all $\alpha$ and that $\Pr\{X_i = X_j\} = 0$ for all $i \neq j$. For $n \geq 2$, define $X_n$ as a record-to-date of the sequence if $X_n > X_i$ for all $i < n$.

   (a) Find the probability that $X_2$ is a record-to-date. Use symmetry to obtain a numerical answer without computation. A one or two line explanation should be adequate.

(b) Find the probability that $X_n$ is a record-to-date, as a function of $n \geq 1$. Again use symmetry.

(c) Find a simple expression for the expected number of records-to-date that occur over the first $m$ trials for any given integer $m$. Hint: Use indicator functions. Show that this expected number is infinite in the limit $m \to \infty$.

**Solution:**

(a) $X_2$ is a record-to-date with probability 1/2. The reason is that $X_1$ and $X_2$ are IID, so either one is larger with probability 1/2; this uses the fact that they are equal with probability 0 since they have a density.

(b) By the same symmetry argument, each $X_i, 1 \leq i \leq n$ is equally likely to be the largest, so that each is largest with probability $1/n$. Since $X_n$ is a record-to-date if and only if it is the largest of $X_1, X_2, \cdots, X_n$, it is a record-to-date with probability $1/n$.

(c) Let $\mathbf{1}_n$ be 1 if $X_n$ is a record-to-date and be 0 otherwise. Thus $\mathbb{E}[\mathbf{1}_n]$ is the expected value of the 'number' of records-to-date (either 1 or 0) on trial $n$. That is

$$\mathbb{E}[\mathbf{1}_n] = \Pr\{\mathbf{1}_n = 1\} = \Pr\{X_n \text{ is a record-to-date}\} = \frac{1}{n}. \tag{40}$$

Thus

$$\mathbb{E}[\text{records-to-date up to } m] = \sum_{n=1}^{m} \mathbb{E}[\mathbf{1}_n] = \sum_{n=1}^{m} \frac{1}{n}. \tag{41}$$

This is the harmonic series, which goes to $\infty$ as $m \to \infty$.

8. (Optional) [Reverse Markov Inequality] Derive the reverse Markov inequality: Let $X$ be a random variable such that $\Pr(X \leq a) = 1$ for some constant $a$. Then for $d < \mathbb{E}X$, we have

$$\Pr(X > d) \geq \frac{\mathbb{E}X - d}{a - d}$$

*Hint: Apply the usual Markov inequality to the new non-negative random variable $a - X$.*

**Solutions:** Define the non-negative random variable $\tilde{X} = a - X$. Then by Markov's inequality,

$$\Pr(X \leq d) = \Pr(a - \tilde{X} \leq d) = \Pr(\tilde{X} \geq a - d) \leq \frac{\mathbb{E}[\tilde{X}]}{a - d} = \frac{a - \mathbb{E}[X]}{a - d}. \tag{42}$$

Thus, by complementation,

$$\Pr(X > d) \geq 1 - \frac{a - \mathbb{E}[X]}{a - d} = \frac{\mathbb{E}[X] - d}{a - d}. \tag{43}$$

9. (Optional) [Knockout Football]

In the knockout phase of a football tournament, there are 32 teams *of equal skill* that compete in an elimination tournament. This proceeds in a number of rounds in which teams compete in pairs; any losing team retires from the tournament. See Fig. 1 for an illustration with 16 teams. What is the probability that two given teams will compete against each other? Generalize your answer to $2^k$ teams.

*The following argument is wrong but the answer is right. There has to be 31 games to knock out all but the ultimate winner. There are $\binom{32}{2}$ possible pairs, so that the probability of a given pair being selected for a particular match is $1/\binom{32}{2} = 1/(16 \cdot 31)$. Since the selection of the teams in the different matches is mutually exclusive, the probability of a given pair being selected is 31 times this, which is $1/16$. Why is this wrong and what's the correct way of doing it?*

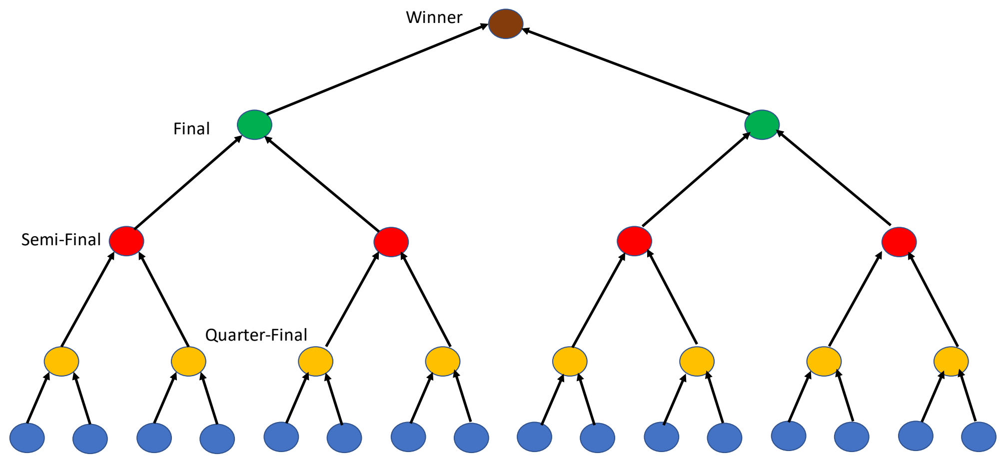


Figure 1: Figure for 16 teams

This problem is taken from Problem 297 of *Five Hundred Mathematical Challenges* (Mathematical Association of America, 1996).

**Solution:** The “solution”' presupposes symmetry among all the matches. The tournament proceeds in several rounds. In the initial round of 16 games, the pairs may indeed be selected at random. But in subsequent rounds, a pair is selected only if both teams survive the previous round. The argument does not use the fact that the teams are of equal abilities. Two teams who are sure to beat everyone else in the tournament will meet with probability 1; two teams inferior to everyone else will meet only in the first round or not at all.

We use induction. After the first round, we have the same situation for 16 players; after the second for 8 players and so on. Let $p_k$ be the probability that a given pair will meet if we start with $2^k$ players of equal ability. Clearly $p_1 = 1$. If we start with 4 players. a given pair will meet in the first round with probability 1/3 and vanquish other players to meet in the second and final round with probability $2/3 \times 1/2 \times 1/2 = 1/6$. Thus $p_2 = 1/2$. In general, we find that

$$p_k = \frac{1}{2^k - 1} + \left[1 - \frac{1}{2^k - 1}\right]\left(\frac{1}{2}\right)^2 p_{k-1}$$

the first term being the probability of the pair meeting in the first round and the second term of each team beating others to get to the second round and eventually meeting. Solving this recursion leads to $p_k = 1/2^{k-1}$. The answer to the problem is indeed 1/16.