Here are solutions to all 27 problems at the end of the notes on *Probability theory* by H.W. Lenstra Jr. (1988). His combinatorial coefficient $\binom{n}{k}$ is here rendered ${}^nC_k = n!/(k!\cdot(n-k)!)$.

1. $$\begin{bmatrix} P(\text{Andrew}) \\ P(\text{Beatrix}) \\ P(\text{Charles}) \end{bmatrix} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 \\ 1 & 1/2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/6 \\ 1/2 \end{bmatrix} .$$

2. Without a Joker, the independent pairs are (A, B) and (A, C) but not (B, C) . With a Joker in the deck, no pair is independent.

3. Let $\mu := P(A)$ and $ß := P(B)$, so that $P(A') = 1-\mu$ and $P(B') = 1-ß$. Since A and B are independent, $P(A\cap B) = \mu\cdot ß$. Then $\mu = P(A) = P(A\cap B) + P(A\cap B') = \mu\cdot ß + P(A\cap B')$, whence follows $P(A\cap B') = \mu\cdot(1-ß)$. Similarly $P(A'\cap B) = (1-\mu)\cdot ß$, and $P(A'\cap B') = (1-\mu)\cdot(1-ß)$. Thus A' and B' are independent because $P(A'\cap B') = P(A')\cdot P(B')$. Likewise for A and B' . But A and A' cannot be independent, since $P(A\cap A') = P(\emptyset) = 0$, unless $\mu = 0$ or $\mu = 1$.

4. Yes, independent because P(divisible by 3)·P(divisible by 5) = (1/3)·(1/5) = 1/15 = = P(divisible by 3 and by 5) . Then P(GCD(number, 15) = 1) = 1 – 1/3 – 1/5 + 1/15 = 8/15 .

5. \$ (5000 + 700 + 3·100)/10000 = \$0.60 .

6. For any number n , the probability that n balls of any particular color will be drawn is the same for every color because of the situation's symmetry: permuting the color's names does not change their probabilities. Therefore the Expected number of balls drawn is the same for every color. Since all three colors' Expected numbers sum to 12 , each Expected number is 4 .

7. (a) $f(x) := x$ **mod** 3 and $g(x) := x$ **mod** 4 , so $x \equiv 4\cdot f(x) + 9\cdot g(x)$ **mod** 12 by the Chinese Remainder Theorem. Since $900 \equiv 0$ **mod** 12 , as s ranges over the set $S := \{1, 2, \ldots, 900\}$ with uniform probability 1/900 per element, $P(f(s) = m) = 1/3$ for each m in {0, 1, 2} and $P(g(s) = n) = 1/4$ for each n in {0, 1, 2, 3} , and then

$$P(f(s) = m \text{ and } g(s) = n) = P(s \equiv 4\cdot m + 9\cdot n \textbf{ mod } 12) = 1/12$$

because $(4\cdot m + 9\cdot n)$ **mod** 12 runs through all 12 members of {0, 1, 2, …, 11} as (m, n) runs through all 12 pairs. Therefore f and g are independent.

7. (b) $E(f+g) = E(f) + E(g) = 1 + 3/2 = 5/2$. $E(f\cdot g) = E(f)\cdot E(g) = 1\cdot 3/2 = 3/2$, and Variance $V(f+g) = V(f) + V(g) = 2/3 + 5/4 = 23/12$ because f and g are independent.

8. E(emales) = 9/3 = 3 ; V(emales) = 9·(2/9) = 2 assuming independence of sex of each birth.

9. The number of ways of choosing 5 volumes out of 10 is ${}^{10}C_5 = 252$. The number of ways to get no complete novel is $2^5 = 32$. The number of ways to get one complete novel is ${}^5C_1\cdot{}^4C_3\cdot 2^3 = 160$. The number of ways to get two complete novels is ${}^5C_2\cdot{}^3C_1\cdot 2^1 = 60$. As a check we observe that 60+160+32 = 252 . Assuming each way as likely as every other,

P(i=0) = 32/252 = 8/63 . P(i=1) = 160/252 = 40/63 . P(i=2) = 60/252 = 15/63 .
E(i) = 70/63 = 10/9 . Variance(i) = 200/567 .

10. This solution takes advantage of three identities obtained by differentiating the first twice :

$$1/(1-q) = \Sigma_{n\ge 0}\, q^n \;; \qquad 1/(1-q)^2 = \Sigma_{n>0}\, n\cdot q^{n-1} \;; \qquad 2/(1-q)^3 = \Sigma_{n>0}\, n\cdot(n-1)\cdot q^{n-2} \;.$$

$$\begin{aligned} V(f) \;:=\; & \Sigma_{n>0}\, q^{n-1}\cdot p\cdot(n - 1/p)^2 \;=\; \Sigma_{n>0}\, q^{n-1}\cdot p\cdot n^2 - 2\cdot\Sigma_{n>0}\, q^{n-1}\cdot n + \Sigma_{n>0}\, q^{n-1}/p \\ =\; & p\cdot q\cdot\Sigma_{n>0}\, n\cdot(n-1)\cdot q^{n-2} + (p-2)\cdot\Sigma_{n>0}\, n\cdot q^{n-1} + \Sigma_{m\ge 0}\, q^m/p \qquad \ldots \text{ recall } (1-q) = p \\ =\; & 2\cdot p\cdot q/p^3 + (p-2)/p^2 + 1/p^2 \;=\; (2\cdot q + p - 2 + 1)/p^2 \;=\; q/p^2 \;. \end{aligned}$$

11. Any positive integer n and nonnegative fraction $p \le 1$ determine a *Binomial Random Variable* f ; it is the count of the successes in n independent Bernoulli trials each with probability p of success : $P(f = k) = {}^nC_k\cdot p^k\cdot(1-p)^{n-k}$. Now

$$P(f \text{ is even}) = \Sigma_{0\le j\le\lfloor n/2\rfloor}\, {}^nC_{2j}\cdot p^{2j}\cdot(1-p)^{n-2j} \quad \text{and} \quad P(f \text{ is odd}) = 1 - P(f \text{ is even}) \;.$$

In the special case that $p = 1/2$ these expressions simplify to $P(f \text{ is even}) = P(f \text{ is odd}) = 1/2$, as is obvious if n is odd because then each term ${}^nC_{2j}\cdot 2^{-n}$ included in the sum $\Sigma\ldots$ can be paired with an equal term ${}^nC_{n-2j}\cdot 2^{-n}$ excluded from that sum; if n is even the simplification of the sum is not obvious. However, the 2^n equally likely outcomes of n trials can be put into pairs that differ only in their first trials; since the members of each pair have opposite even-odd parity, the odd counts f must be as numerous as the even counts and equally likely.

12. The same reasoning as worked in Example 7 implies that 2 is the expected number of children whose first name starts with the letter that they get. The variance cannot be determined because it is positive unless all children's names begin with the same letter, in which case the variance is zero.

13. There are $2^{100} \approx 1.26765\cdot 10^{30}$ subsets of a set with 100 elements. The subsets with cardinalities between 41 and 59 inclusive number

$$\Sigma_{41\le k\le 59}\, {}^{100}C_k = {}^{100}C_{50} + 2\cdot\Sigma_{1\le k\le 9}\, {}^{100}C_{50-k} \approx 1.19554\cdot 10^{30} \;.$$

Their ratio is $1.19554/1.26765 \approx 0.9431$. This ratio is the same as the probability that a binomial random variable counting the successes in $n = 100$ independent Bernoulli trials with $p = 1/2$ will depart from the mean $n\cdot p = 50$ by less than $10/5 = 2$ times the standard deviation $\sqrt{(n\cdot p\cdot(1-p))} = 5$. Chebyshev's Inequality says this probability is at least $1 - 1/2^2 = 0.75$. The Central Limit Theorem uses a *Normal* random variable (see the class notes titled *Law of Large Numbers*) distributed continuously with the same mean and standard deviation to estimate that a departure from the mean smaller than μ times the standard deviation has probability near $\Phi(\mu) - \Phi(-\mu)$; but what value of μ should be used for the given *discrete* distribution?

If $\mu := 10/5 = 2$	then $\Phi(\mu) - \Phi(-\mu) \approx 0.955$	overestimates the probability.
If $\mu := 9/5 = 1.8$	then $\Phi(\mu) - \Phi(-\mu) \approx 0.928$	underestimates the probability.
If $\mu := 9.5/5 = 1.9$	then $\Phi(\mu) - \Phi(-\mu) \approx 0.943$	gets it about right.

14. Random variable f has mean $\bar{f} := E(f)$ and variance $\sigma^2 := E((f-\bar{f})^2)$, and random variable $g := |f - \bar{f}|$ has mean $\tau := E(g)$. The variance of g is

$$V := E((g-\tau)^2) = E(g^2) - 2\cdot\tau\cdot E(g) + \tau^2 = E((f-\bar{f})^2) - 2\cdot\tau^2 + \tau^2 = \sigma^2 - \tau^2 .$$

Since all three of V , σ and τ are nonnegative, $\sigma \ge \tau$ as claimed, with equality just when $V = 0$, which occurs just when g is constant (τ) instead of random, which occurs just when f takes at most two values each with probability $1/2$.

15. Let random variable f have mean $\bar{f} := E(f)$ and variance $\sigma^2 := E((f-\bar{f})^2)$. Chebyshev's Inequality says $P(|f-\bar{f}| \ge r) \le \sigma^2/r^2$ (meaningfully only if $r \ge \sigma$) because

$$\sigma^2 = \sum_{\text{all } t} (t-\bar{f})^2\cdot P(f=t) = \sum_{|t-\bar{f}|<r} (t-\bar{f})^2\cdot P(f=t) + \sum_{|t-\bar{f}|\ge r} (t-\bar{f})^2\cdot P(f=t)$$
$$\ge \sum_{|t-\bar{f}|\ge r} (t-\bar{f})^2\cdot P(f=t) \ge \sum_{|t-\bar{f}|\ge r} r^2\cdot P(f=t) = r^2\cdot P(|f-\bar{f}| \ge r) ,$$

with equality just when $\sum_{|t-\bar{f}|<r} (t-\bar{f})^2\cdot P(f=t) = 0$ and $\sum_{|t-\bar{f}|>r} r^2\cdot P(f=t) = 0$. The first equation here implies $P(0 < |f-\bar{f}| < r) = 0$ and the second implies $P(|f-\bar{f}| > r) = 0$, so random variable f can take on at most three values $\bar{f}-r$, $\bar{f}$ and $\bar{f}+r$. Apparently $P(f = \bar{f}-r) = P(f = \bar{f}+r)$ to ensure that $E(f) = \bar{f}$; and to ensure that $E((f-\bar{f})^2) = \sigma^2$ we find $P(f = \bar{f}-r) = P(f = \bar{f}+r) = \sigma^2/(2\cdot r^2)$. Then $P(f = \bar{f}) = 1 - \sigma^2/r^2$ and Chebyshev's Inequality becomes an equality, barely.

16. Random variable f takes the value 1 with probability $p := P(A)$, 0 with probability $(1-p) = P(A')$. Then mean $\bar{f} := E(f) = p\cdot 1 + (1-p)\cdot 0 = p$, and the variance of f is

$$p\cdot(1-p)^2 + (1-p)\cdot(0-p)^2 = p\cdot(1-p) \quad \text{as claimed.}$$

17. For each ordering in which the list's i-th item is marked, we obtain $i-1$ equally likely orderings in which that i-th item is unmarked by swapping the i-th item with one of its $i-1$ lesser predecessors; and all orderings are enumerated once apiece this way. Thus the probability that the i-th item is marked is $1/i$. Therefore the expected number of marked items is

$$H_{20} = 1 + 1/2 + 1/3 + \ldots + 1/20 \approx 3.59774 .$$

Had the number of items been a number N rather bigger than 20 , the *Harmonic Number* H_N would have been better approximated as described in the class notes on *Some Inequalities*.

18. See the class notes about *Derangements*, or our text, for the facts about the number D_n of derangements of a set of n elements; this is the number of permutations that leave no member of the set unmoved. There we find that $D_0 = 1$, $D_1 = 0$, $D_2 = 1$, and for $n = 1, 2, 3, \ldots$ in turn

$$D_n = n\cdot D_{n-1} + (-1)^n .$$

18(a). P(no one gets his hat back) $= D_{10}/10! = (10\cdot D_9 + 1)/10!$
$> (10\cdot D_9)/10! =$ P(just one man gets his hat back) .

If the number of men were N instead of 10 , the inequality would go the same way for all even $N > 1$, the opposite way for all odd $N > 0$.

18(b). P(Just nine out of ten men get their hats back) = 0 ; it can't happen.

18(c). For any i chosen in advance from the set {1, 2, …, 10} , Example 7 explained why Probability(Man #i gets his hat back) = 1/10 . For any two *different* i and j chosen in advance from that set, Probability(Men #i and #j get their hats back) = (10–2)**!**/10**!** = 1/90 . This differs from 1/100 , which is what this probability would be if the events "Man #i gets his hat back" and "Man #j gets his hat back" were independent. These events are correlated; when one occurs it enhances the other's likelihood.

18(d). Let f_i := (1 if man #i gets his hat back, 0 otherwise) , and let $f := \sum_i f_i$. We already know from Example 7 that Probability($f_i = 1$) = 1/10 , so that $E(f_i) = 1/10$ and therefore f 's expected value $\overline{f} := E(f) = 1$; and there is a formula for f 's variance: $V := E(f^2) - \overline{f}^2$. To compute this we need to know that $E(f_i^2) = E(f_i) = 1/10$ and, from 18(c), that $E(f_i \cdot f_j) = 1/90$ if $i \neq j$. Then $E(f^2) = E((\sum_i f_i)^2) = \sum_i E(f_i^2) + \sum_i \sum_{j\neq i} E(f_i \cdot f_j) = 10/10 + 90/90 = 2$, whence variance $V = 2 - 1 = 1$.

19. Let $f_k := P(f = k)$ and $a_k := P(f \geq k) = \sum_{j\geq k} f_j$ for every integer $k \geq 0$; in particular $a_0 = 1$ because we are told that f takes only nonnegative integer values. Then we find that

$$\sum_{k\geq 1} a_k = \sum_{k\geq 1} \sum_{j\geq k} f_j = \sum_{j\geq 1} \sum_{1\leq k\leq j} f_j \quad \text{… after the order of summation is reversed}$$
$$= \sum_{j\geq 1} j \cdot f_j = E(f) .$$

20. This problem is essentially the same as the *Monty Hall Three Door Puzzle*, Example 10 on p. 265 of our text *Discrete Mathematics and its Applications* 4th. ed. by K.H. Rosen (1999, McGraw-Hill). In terms of coin and cups, I have two choices only one of which is random. The random choice answers the question "To which of three cups will I point first?". The non-random choice selects in advance one of two strategies:

<u>Strategy 1</u> : Uncover the cup to which I first pointed. Its probability of covering my coin is 1/3 , which is not changed by the performer's knowledge (which I do not share) of which of the other two cups does *not* cover my coin. When he lifts that cup he does not change what I shall receive.

<u>Strategy 2</u> : Uncover not the cup to which I first pointed but the other cup still unlifted. The probability that the cup to which I first point does *not* cover my coin is 2/3 . This probability is not changed by the performer's revelation of the other cup not covering my coin when he lifts it, so 2/3 is also the probability that my coin will lie under the third cup, the cup I will uncover.

Evidently strategy 2 is twice as likely as strategy 1 to recover my coin. The sum of their probabilities, 1/3 + 2/3 = 1 , corresponds to a third strategy of the kind favored by historical figures like Alexander the Great at Gordium: knock over both unlifted cups.

21. This hard exercise challenges the student to construct an accurate mathematical model for a complicated game. It is complicated partly because it can terminate in two ways, either in step (i) or in step (ii), and each way requires its own peculiar proof. The first proof treats values of the "debt" x that compell the game to terminate either in step (ii) or in step (i) after at most a predetermined finite number k of coin-flips. The second proof will treat "debts" x that prevent the game from terminating in step (i), allowing it to continue for arbitrarily many coin-flips (though with ever diminishing probability) until it terminates in step (ii).

If the holder of the coin (initially Klaas) owe a "debt" x to the non-holder (initially Karel), let $æ(x)$ be the non-holder's expectation, so $1–æ(x)$ is the coin-holder's expectation. Evidently $æ(0) = 0$ and $æ(1) = 1$, both determined by step (i) of the game without any coin-flip. When $0 < x \le 1/2$ the coin-holder can expect to win the coin in step (ii) with probability $1/2$ and end the game leaving nothing for the non-holder, or else with probability $1/2$ the "debt" will double and the game continue with the same coin-holder; consequently

$$\text{if } 0 \le x \le 1/2 \quad \text{then} \quad æ(x) = 0/2 + æ(2x)/2 = æ(2x)/2 \qquad (\dagger)$$

is the non-holder's expectation, and the coin-holder's expectation is $1–æ(2x)/2$. On the other hand when $1/2 < x < 1$ step (i) reverses the roles of coin-holder and non-holder and replaces the "debt" x by $1–x$; consequently

$$\text{if } 1/2 < x \le 1 \quad \text{then} \quad æ(x) = 1 – æ(2(1–x))/2 \qquad (\ddagger)$$

is the former non-holder's expectation. The two equations (†) and (‡) are *functional equations* from which we shall deduce by mathematical induction that $æ(x) = x$ whenever $0 \le x \le 1$, but our process of deduction will work on only those values x for which the game can't run forever.

The game can terminate in step (i) only if x is an integer multiple of a power of $1/2$, which case will be considered now. Consider $x = m/2^k$ for nonnegative integers k and $m \le 2^k$ (since $0 \le x \le 1$). Every coin-flip that shows tails will replace x by either $2x$ or $2(1–x)$; both replacements will decrement k by 1 , so at most k coin-flips can occur, after which the game must terminate in step (i). For just such values $x = m/2^k$ we shall use induction on k to prove the formula $æ(m/2^k) = m/2^k$. We have already noted its validity when $k = 0$ and $m = 0$ and $m = 1 = 2^0$. Let our induction hypothesis be that the formula is valid for some $k = K \ge 0$ while $0 \le m \le 2^K$. To verify the formula for $k = K+1$ we need examine just two cases:

- If $0 \le m \le 2^K$ then $0 \le m/2^{K+1} \le 1/2$ and equation (†) plus the induction hypothesis implies
$$æ(m/2^{K+1}) = æ(m/2^K)/2 = m/2^{K+1} \text{ as claimed.}$$
- If $2^K < m \le 2^{K+1}$ then $1/2 < m/2^{K+1} \le 1$ and equation (‡) instead of (†) implies
$$æ(m/2^{K+1}) = 1 – æ((2^{K+1}–m)/2^K)/2 = m/2^{K+1} \quad \text{because } 0 \le 2^{K+1}–m < 2^K .$$

Therefore $æ(x) = x$ whenever x is an integer multiple of a power of $1/2$ between 0 and 1 .

We could infer that Karel's expectation $æ(x) = x$ for *all* real values x between 0 and 1 if we proved that $æ(x)$ is a monotone increasing function of x . Such a proof is feasible but would suit students of Real Analysis (Math. 104) better than students of Discrete Math. (55); besides it would reproduce a large fraction of the next proof, which runs along lines that Lenstra seems to have intended, judging by his hint. However the next proof works for only those "debts" x that let the game run forever albeit with probability 0 .

First let us model coin-tosses as a sample-space of infinitely many mutually exclusive outcomes H, TH, TTH, …, $T^{k-1}H$, … in which " $T^{k-1}H$ " stands for k coin-flips of which the first k–1 show tails and the last shows the head of Queen Beatrix. The probability of $T^{k-1}H$ is $1/2^k$ because the coin is fair. Note that all these probabilities add up to $\sum_{k>0} 1/2^k = 1$, leaving zero for the probability that the game will run forever because the Queen's head never shows up. (A Dutch guilder with no head is unfair and illegal, so Klaas can't possibly have one.)

To determine who will hold the coin for the k-th flip, let us represent the "debt" x as a *twos-complement* binary fraction $x = -x_0 + \sum_{j>0} x_j/2^j$ in which each bit x_j is either 0 or 1 , but we disallow the possibility that all but finitely many of those bits are the same. In other words, we allow no binary representation to end with an infinite string of ones nor with an infinite string of zeros, thus excluding values x that are integer multiple of a power of 1/2 ; this exclusion is tolerable because such values have already been handled by the previous proof. Except for such values, every other number x in the range $-1 < x < 1$ has a nontrivially nonterminating twos-complement binary representation, as the reader should be able to verify, although the initial "debt" x lies in the narrower range $0 < x < 1$. For example, $1/3 = \texttt{0.01010101}\ldots$ in binary and $-1/3 = \bar{\texttt{1}}\texttt{.10101010}\ldots = -1 + 2/3$. More generally, $x > 0$ when $x_0 = 0$ and $x < 0$ when $x_0 = 1$; in other words, $\mathrm{Sign}(x) := 1-2x_0 = \pm 1$ as expected. (Sign(0) won't occur.)

Because x cannot be an integer multiple of a power of 1/2 , the same is true for its replacements 2x and 2(1–x) in the course of the game. Therefore the game cannot terminate in step (i) ; only the appearance of the Queen's head in step (ii) can end the game, and then the coin-holder who flipped the coin will retain it. Our next task is to figure out how the bits of x determine who that coin-holder will be.

To that end set $X_0 := x$, so $0 < X_0 < 1$, and for k = 1, 2, 3, … in turn define X_k thus:

If $|X_{k-1}| < 1/2$ then set $X_k := 2\cdot X_{k-1}$; otherwise, since $X_{k-1} \neq \pm 1/2$,

when $1/2 < |X_{k-1}| < 1$, set $X_k := 2\cdot\big(X_{k-1} - \mathrm{Sign}(X_{k-1}) \big)$.

Either way, mathematical induction confirms that $-1 < X_k = -x_k + \sum_{j>k} x_j/2^{j-k} < 1$ for all $k > 0$ as follows: The tested condition " $|X_{k-1}| < 1/2$ " is tantamount to " $x_{k-1} = x_k$ " and, if true, ensures that $|X_k| < 1$ and $\mathrm{Sign}(X_k) := 1-2x_k = \mathrm{Sign}(X_{k-1})$. (*Overflow* could spoil these last equations only if x were an odd integer multiple of $1/2^k$, but this possibility has been ruled out.) If the alternative condition " $|X_{k-1}| > 1/2$ " tantamount to " $x_{k-1} \neq x_k$ " is true it ensures that $0 < |X_k| < 1$ and $\mathrm{Sign}(X_k) = -\mathrm{Sign}(X_{k-1})$. (*Overflow* could spoil this last equation only if x were an odd integer multiple of $1/2^k$.) Apparently $x = |X_k/2|$ is the "debt" owed by the coin-holder to the non-holder immediately before the k-th flip, and $\mathrm{Sign}(X_k)$ reverses just when the coin changes hands in step (i). Immediately after the k-th coin flip, if it shows tails, X_k is the signed "debt" Klaas owes Karel, positive if Klaas owes $|X_k|$ to Karel, negative if Karel owes $|X_k|$ to Klaas. Here is a summary of how the signed "debt" X_k correlates with the game:

(o) Initially Klaas, who holds the coin, owes Karel a "debt" of $x = X_0$, where $0 < x < 1$.

(i) After $k-1$ coin-flips, the "debt" stands at $x = |X_{k-1}|$; if $X_{k-1} > 0$ Klaas owes it to Karel and holds the coin, but if $X_{k-1} < 0$ then Karel owes x to Klaas and holds the coin. If now $x < 1/2$, set $X_k := 2\cdot X_{k-1}$ so that $x = |X_k|/2$. Otherwise, when $1/2 < x < 1$, the coin must change hands and the "debt" change from x to $1-x = |X_k|/2$ in the opposite direction, where now $\mathrm{Sign}(X_k) = -\mathrm{Sign}(X_{k-1})$. The signed "debt" $X_k/2$ must still be nonzero because it is never an integer multiple of a power of $1/2$.

(ii) The coin-holder is now Klaas if $x_k = 0$, Karel if $x_k = 1$. The coin-holder, who owes the non-holder $|X_k|/2$, flips the coin. If the queen's head shows after this k-th flip, the coin-holder retains the coin and ends the game. Otherwise the debt is doubled, k is incremented by 1 , and the game continues from step (i) above.

If the k-th coin-flip ends the game, an outcome $T^{k-1}H$ with probability $1/2^k$, the game ends with Klaas holding the coin if $x_k = 0$, Karel if $x_k = 1$. Therefore Karel's expectation is

$$æ(x) := \sum_{k>0} x_k/2^k = x ,$$

as claimed.

22. The probability generating function of the random variable f is $g(x) := \sum_{k\ge 0} f_k\cdot x^k$ wherein $f_k = P(f = k) \ge 0$. Since f takes only nonnegative integer values, $g(1) = \sum_{k\ge 0} f_k = 1$. The assumption that " the radius of convergence of g is larger than 1 " is essential because otherwise, for instance if $f_k = 1/((k+1)\cdot(k+2))$, the expected value $E(f) = g'(1)$ could be ∞ . Thus we can assume that the expected value $\overline{f} := E(f) = \sum_{k\ge 0} k\cdot f_k = g'(1)$ is finite. The variance

$$\begin{aligned} V(f) = E((f-\overline{f})2) = E(f^2) - \overline{f}^2 &= \sum_{k\ge 0} k^2\cdot f_k - g'(1)^2 \\ &= \sum_{k\ge 0} k\cdot(k-1)\cdot f_k + \sum_{k\ge 0} k\cdot f_k - g'(1)^2 \\ &= g''(1) + g'(1) - g'(1)^2 . \end{aligned}$$

23. When Lenstra wrote "… maps the n-th variable to n " I think he meant "… maps the n-th element of {h, mh, mmh, mmmh, …} to n." Thus, $P(f = n) = q^{n-1}\cdot p$. The probability generating function of f is $g(x) := \sum_{n\ge 1} q^{n-1}\cdot p\cdot x^n = p\cdot x/(1 - q\cdot x)$. The expected value $\overline{f} := E(f) = g'(1) = p/(1-q) + p\cdot q/(1-q)^2 = 1/p$ because $1-q = p$. The variance of f is

$$V = g''(1) + g'(1) - g'(1)^2 = 2p\cdot q/(1-q)^2 + 2p\cdot q^2/(1-q)^3 + 1/p - 1/p^2 = q/p^2 .$$

24. Lenstra's notation for this problem has too many subscripts, so let's simplify it. Let f, g and $h := f + g$ be random variables that take exclusively nonnegative integer values, and let their respective probability generating functions be $F(x) := \sum_{k\ge 0} f_k\cdot x^k$, $G(x) := \sum_{k\ge 0} g_k\cdot x^k$, and $H(x) := \sum_{k\ge 0} h_k\cdot x^k$. Provided f and g are *independent*, …

$$\begin{aligned} h_k = P(h = k) &= P(f + g = k) = \textstyle\sum_{0\le j\le k} P(f = j \ \&\ g = k-j) \\ &= \textstyle\sum_{0\le j\le k} P(f = j)\cdot P(g = k-j) \qquad \ldots \text{ because } f \text{ and } g \text{ are independent} \\ &= \textstyle\sum_{0\le j\le k} f_j\cdot g_{k-j} \ , \end{aligned}$$

which is the same as the coefficient of x^k in

$$F(x)\cdot G(x) = (\textstyle\sum_{m\ge 0} f_m\cdot x^m)\cdot(\sum_{n\ge 0} g_n\cdot x^n) = \sum_{m\ge 0}\sum_{n\ge 0} f_m\cdot g_n\cdot x^{m+n} = \sum_{k\ge 0} x^k\cdot\sum_{0\le j\le k} f_j\cdot g_{k-j} \ .$$

Therefore $H(x) = F(x)\cdot G(x)$, though to believe this you may have to take for granted infinite series manipulations whose validity is established in advanced Calculus courses.

25. As in Exercise 17, the sample space S is the set of all lists formed by permuting a given set of N distinct numbers. ($N = 20$.) Each such list is as likely as any other to be chosen at random. The first element in each list is marked and so is its every element that exceeds all its predecessors in the list, but no other elements are marked. For any positive integer $i \le N$ let A_i denote the subset of lists whose i-th element is among the marked elements. Exercise 17's solution explained why $P(A_i) = \#(A_i)/\#(S) = 1/i$. Note that every $A_i\cap A_j$ is nonempty.

25(a) For every positive integer $K \le N$ and for every subset of K integers $\{i, j, \ldots, k\}$ drawn from the set $\{1, 2, 3, \ldots, N\}$ we wish to show that $P(A_i\cap A_j\cap\ldots\cap A_k) = P(A_i)\cdot P(A_j)\cdot\ldots\cdot P(A_k)$. The formula is trivially true for $K = 1$; and the formula is true for $K = N$ because only one list lies in $A_1\cap A_2\cap\ldots\cap A_N$ so its probability $P(A_1\cap A_2\cap\ldots\cap A_N) = 1/N! = P(A_1)\cdot P(A_2)\cdot\ldots\cdot P(A_N)$ thanks to Exercise 17 . Let's prove the formula for every K between 1 and N by induction: Suppose the formula is true for some K between 1 and $N-1$ inclusive. What about $K+1$? Let the subset of $K+1$ integers be $\{i, j, \ldots, k, m\}$; there is no loss of generality in assuming them so ordered that $i < j < \ldots < k < m$ since doing so changes neither $A_i\cap A_j\cap\ldots\cap A_k\cap A_m$ nor $P(A_i)\cdot P(A_j)\cdot\ldots\cdot P(A_k)\cdot P(A_m)$. The lists in $A_j\cap\ldots\cap A_k\cap A_m$ can now be partitioned into two subsets: the lists in $A_i\cap A_j\cap\ldots\cap A_k\cap A_m$ and the rest in $A_j\cap\ldots\cap A_k\cap A_m - A_i$. From each list in $A_i\cap A_j\cap\ldots\cap A_k\cap A_m$ we may generate $i-1$ lists in $A_j\cap\ldots\cap A_k\cap A_m - A_i$ by swapping the marked i-th item in the former list with one of its $i-1$ lesser predecessors; doing so generates every list in $A_j\cap\ldots\cap A_k\cap A_m$ just once, so $\#(A_j\cap\ldots\cap A_k\cap A_m) = i\cdot\#(A_i\cap A_j\cap\ldots\cap A_k\cap A_m)$. This implies $P(A_i\cap A_j\cap\ldots\cap A_k\cap A_m) = P(A_j\cap\ldots\cap A_k\cap A_m)/i = P(A_i)\cdot P(A_j\cap\ldots\cap A_k\cap A_m)$ since all the lists are equally likely. Finally invoke the induction hypothesis upon the K integers $\{j, \ldots, k, m\}$ to infer $P(A_i\cap A_j\cap\ldots\cap A_k\cap A_m) = P(A_i)\cdot P(A_j)\cdot\ldots\cdot P(A_k)\cdot P(A_m)$ as claimed.

25(b) The number of marked elements in a list is a random variable $f := f_1 + f_2 + \ldots + f_N$ where $f_i = 1$ just when the randomly selected list's i-th element exceeds all its predecessors in the list; otherwise $f_i = 0$. In other words, $f_i = 1$ just for the lists that lie in A_i , so $P(f_i = 1) = 1/i$. Moreover, we have just seen in 25(a) that the random variables f_i are independent. Therefore variance $V(f) = V(f_1) + V(f_2) + \ldots + V(f_N)$. Now, $f_i^2 = f_i = 0$ or 1 , $E(f_i) = P(f_i = 1) = 1/i$ and $V(f_i) = E(f_i^2) - E(f_i)^2 = 1/i - 1/i^2$, so

$$\begin{aligned} V(f) &= (\ 1 + 1/2 + 1/3 + \ldots + 1/N\ ) - (\ 1 + 1/4 + 1/9 + \ldots + 1/N^2\ ) \\ &\approx 2.001576 \ \text{ when } N = 20 \ . \end{aligned}$$

25(c) Exercise 24 showed that the probability generating function of a sum of independent nonnegative integer-valued random variables is the product of their individual probability generating functions. The probability generating function for f_i is $(1–1/i) + z/i$, so the probability generating function for their sum f must be $z\cdot(1+z)\cdot(2+z)\cdot(\ldots)\cdot(N–1+z)/N!$, as claimed.

26(a). This problem resembles Lenstra's Example 8. After buying one box, the probability that any subsequent purchase will contain a letter different from the first's is $p = 5/6$, analogous to the probability of hitting a target. The expected number of additional purchased boxes, like the expected number of shots until the first hit, is $1/p = 6/5$. Therefore the total number of purchased boxes expected before two different letters are acquired is $1 + 6/5$.

26(b). Let p_j be the probability that any purchased box will contain a letter different from j previously chosen letters. Evidently $p_j = 1 – j/6$. Let random variable n_j be the number of boxes purchased to get a letter different from j previously chosen letters. As before, we find $E(n_j) = 1/p_j = 6/(6–j)$. No claim is made yet that these random variables n_j are independent. Still, the number of boxes purchased to get at least one instance of every letter is a random variable $n := \sum_{0\le j\le 5} n_j$ because, whatever the order in which letters appear, the sequence of purchases can be broken into batches, each a batch of purchases of which the last acquired a letter not seen before. The problem's solution is $E(n) = \sum_{0\le j\le 5} E(n_j) = \sum_{0\le j\le 5} 6/(6–j) = 147/10$.

26(c). As in Exercise 23, when $0 \le j \le 5$ we find the probability generating function for n_j to be $N_j(x) = \sum_{k\ge 1} p_j\cdot(1–p_j)^{k–1}\cdot x^k = (1–j/6)\cdot x/(1 – j\cdot x/6)$. Moreover each n_j is independent of the others because, first, it is unchanged by changes among the prior set of j letters in boxes already purchased, and second because n_j does not depend upon which new letter turns up so long as it is different from the prior set of j. Thanks to Exercise 24, the probability generating function for $n = \sum_{0\le j\le 5} n_j$ is $N(x) = \prod_{0\le j\le 5} N_j(x) = (5!/6^5)\cdot x^6/\prod_{1\le j\le 5} (1 – j\cdot x/6)$.

27. See the class notes on *Derangements* to find out about D_n, the number of derangements of n objects (permutations that leave no object unmoved), and about the number $F_{n,k} = {}^nC_k\cdot D_{n–k}$ of permutations of n objects that leave exactly k unspecified objects unmoved. Those notes explain why $D_n = n\cdot D_{n–1} + (–1)^n = n!\cdot\sum_{0\le j\le n} (–1)^j/j!$.

27(a). $P(f_n = k) = F_{n,k}/n! = (1/k!)\cdot\sum_{0\le j\le n–k} (–1)^j/j!$ for $0 \le k \le n$.

27(b).
$$\begin{aligned} g_n(z) &= \sum_{k\ge 0} (z^k/k!)\cdot\sum_{0\le j\le n–k} (–1)^j/j! = \sum_{0\le k\le n} \sum_{0\le j\le n–k} (z^k/k!)\cdot(–1)^j/j! \ \ldots \text{ set } m := j+k \\ &= \sum_{0\le m\le n} \sum_{0\le k\le m} (z^k/k!)\cdot(–1)^{m–k}/(m–k)! = \sum_{0\le m\le n} (z–1)^m/m!, \text{ so} \end{aligned}$$
$$\begin{aligned} (1–t)\cdot\sum_{n\ge 0} g_n(z)\cdot t^n &= \sum_{n\ge 0} \sum_{0\le m\le n} (t^n–t^{n+1})\cdot(z–1)^m/m! = \sum_{m\ge 0} \sum_{n\ge m} (t^n–t^{n+1})\cdot(z–1)^m/m! \\ &= \sum_{m\ge 0} t^m\cdot(z–1)^m/m! \ \ldots \text{ because } \sum_{n\ge m} (t^n–t^{n+1}) \text{ telescopes} \\ &= \exp(t\cdot(z–1)) \text{ as claimed.} \end{aligned}$$