CS 70 Discrete Mathematics and Probability Theory
Fall 2012 Vazirani HW 13

1 Variance

1. Let X and Y be independent random variables. Express $Var(X - Y)$ in terms of $Var(X)$ and $Var(Y)$.

–Solution–

$$\mathbf{Var}[X - Y] = \mathbf{Var}[X + (-Y)]$$

X and Y are independent, and therefore so are X and $-Y$.

$$= \mathbf{Var}[X] + \mathbf{Var}[-Y]$$
$$= \mathbf{Var}[X] + \mathbf{Var}[(-1) \cdot Y]$$

We've proved in the previous homework that $\mathbf{Var}[cX] = c^2\mathbf{Var}[X]$.

$$= \mathbf{Var}[X] + (-1)^2\mathbf{Var}[Y]$$
$$= \mathbf{Var}[X] + \mathbf{Var}[Y]$$

X and Y are independent, so we have

$$= \mathbf{Var}[X + Y]$$

2. Given a random variable X with $E(X) = \mu$ and $Var(X) = \sigma^2$, let the random variable $Y = XZ$ where Z is defined as follows: Flip a fair coin. If it comes up H, then $Z = 1$, otherwise $Z = -1$. Find $E[Y]$ and $Var(Y)$.

–Solution–

$$\mathbf{E}[Y] = \mathbf{E}[XZ]$$

X and Z are independent

$$= \mathbf{E}[X]\,\mathbf{E}[Z]$$

We know that $\mathbf{E}[Z] = 0$

$$= 0$$

$$\mathbf{E}\left[Y^2\right] = \mathbf{E}\left[X^2 Z^2\right]$$

X and Z are independent, and therefore so do X^2 and Y^2

$$= \mathbf{E}\left[X^2\right]\mathbf{E}\left[Z^2\right]$$

Z^2 is always 1. Therefore, $\mathbf{E}\left[Z^2\right] = 1$

$$= \mathbf{E}\left[X^2\right]$$

We know that $\mathbf{Var}[X] = \mathbf{E}\left[X^2\right] - \mathbf{E}[X]^2$, and therefore

$$= \sigma^2 + \mu^2$$

$$\begin{aligned}\mathbf{Var}[Y] &= \mathbf{E}\left[Y^2\right] - \mathbf{E}[Y]^2 \\ &= \mathbf{E}\left[Y^2\right] - 0 \\ &= \sigma^2 + \mu^2\end{aligned}$$

2 Coupon Collecting

1. Let X be the number of tosses of a biased coin with Heads probability p until the first Head appears (i.e. X is a geometric r.v. with parameter p). We have seen in lecture that $\mathbf{E}[X] = \frac{1}{p}$. Show that $\mathbf{Var}[X] = \frac{1-p}{p^2}$.
   Hint: You will need to sum a series like $S = \sum_{i=1}^{\infty} i^2 q^i$. One way to do this is to multiply S by q and subtract the result from S: this gives you a series for $(1-q)S$. Now if you look at this series carefully, you will see that you can split it into a series of the form $\sum_i iq^i$ and one of the form $\sum_i q^i$. But you know how to sum both of these: the first is like the expectation of X and the second is just a geometric series.)

–Solution–

$$\mathbf{E}\left[X^2\right] = \sum_{k=1}^{\infty} k^2(1-p)^{k-1}p$$

Take $q = 1 - p$

$$\begin{aligned}
&= \sum_{k=1}^{\infty} k^2 q^{k-1}(1-q) \\
&= \frac{1}{q}(1-q)\sum_{k=1}^{\infty} k^2 q^k \\
&= \frac{1}{q}(1-q)S
\end{aligned}$$

We can now analyze the hint a very small step at a time

$$\begin{aligned}
S(1-q) &= S - qS \\
&= \sum_{k=1}^{\infty} k^2 q^k - \sum_{k=1}^{\infty} k^2 q^{k+1}
\end{aligned}$$

We can now change the summation bound from 1 to 2 in the second expression, compensating by changing k to $k-1$. This does not change the result.

$$= \sum_{k=1}^{\infty} 1k^2 q^k - \sum_{k=2}^{\infty} (k-1)^2 q^k$$

with $k = 1$, the expressoin $(k-1)q^k = 0$, so we can change the summation bound to 1 without changing the result

$$\begin{aligned}
&= \sum_{k=1}^{\infty} k^2 q^k - \sum_{k=1}^{\infty} (k-1)^2 q^k \\
&= \sum_{k=1}^{\infty} q^k (k^2 - (k-1)^2) \\
&= \sum_{k=1}^{\infty} q^k (k^2 - k^2 + 2k - 1) \\
&= \sum_{k=1}^{\infty} q^k (2k-1) \\
&= 2\sum_{k=1}^{\infty} k q^k - \sum_{k=1}^{\infty} q^k
\end{aligned}$$

Take $p = 1 - q$

$$= 2\sum_{k=1}^{\infty} k(1-p)^k - \sum_{k=1}^{\infty}(1-p)^k$$

Manipulate a bit to get results that we are already familiar with

$$= 2\frac{1-p}{p}\sum_{k=1}^{\infty} k(1-p)^{k-1}p - \sum_{k=1}^{\infty}(1-p)^k$$

The first sum is the expected value of a geometric r.v. with parameter p and the second one is a geometric series

$$= 2\frac{1-p}{p}\cdot\frac{1}{p} - \frac{1-p}{p}$$
$$= 2\frac{1-p}{p^2} - \frac{1-p}{p}$$

We can now conclude

$$\begin{aligned}\mathbf{Var}[X] &= \mathbf{E}\left[X^2\right] - \mathbf{E}\left[X^2\right] \\ &= \frac{1}{1-p}\left(2\frac{1-p}{p^2} - \frac{1-p}{p}\right) - \frac{1}{p^2} \\ &= \frac{2}{p^2} - \frac{p}{p^2} - \frac{1}{p^2} \\ &= \frac{1-p}{p^2}\end{aligned}$$

An alternative solution, without using the hint

–Solution–

$$\begin{aligned}\mathbf{E}\left[X^2\right] &= \sum_{k=1}^{\infty} k^2 p(1-p)^{k-1} \\ &= p + \sum_{k=2}^{\infty} k^2 p(1-p)^{k-1}\end{aligned}$$

Let $v = k - 1$, so

$$\begin{aligned}\mathbf{E}\left[X^2\right] &= p + (1-p)\sum_{v=1}^{\infty}(v+1)^2 p(1-p)^{v-1} \\ &= p + (1-p)\left[\sum_{v=1}^{\infty} v^2 p(1-p)^{v-1} + 2\sum_{v=1}^{\infty} vp(1-p)^{v-1} + \sum_{v=1}^{\infty} p(1-p)^{n-1}\right] \\ &= p + (1-p)\left[\mathbf{E}\left[X^2\right] + 2\mathbf{E}\left[X\right] + 1\right]\end{aligned}$$

We therefore have

$$\begin{aligned}p\mathbf{E}\left[X^2\right] &= p + \frac{2(1-p)}{p} + (1-p) \\ \mathbf{E}\left[X^2\right] &= \frac{2-p}{p^2}\end{aligned}$$

To conclude

$$\begin{aligned}\mathbf{Var}\left[X^2\right] &= \mathbf{E}\left[X^2\right] - \mathbf{E}\left[X\right]^2 \\ &= \frac{2-p}{p^2} - \frac{1}{p^2} \\ &= \frac{1-p}{p^2}\end{aligned}$$

2. Now let X be the r.v. in the coupon collecting problem, i.e. X is the number of cereal boxes we need to buy before we have collected one copy of each of n baseball cards. Recall from lecture that $\mu = \mathbf{E}[X] = n\sum_{i=1}^{n}\frac{1}{i} \approx n(\ln n + \gamma)$ for a constant $\gamma \approx 0.5$. Use the result of part (a) to compute the variance $\mathbf{Var}[X]$. [Note: your answer should contain a sum of the form $\sum_{i=1}^{n}\frac{1}{i^2}$.]

–Solution–

Recall that X the r.v. from the coupon collecting problem is described by

$$X = \sum_{k=1}^{n} X_k,$$

where X_i is a geometric random variable with probability of success

$$p_k = \frac{n-k+1}{n}$$

Since the X_k-s are all independent,

$$\mathbf{Var}[X] = \sum_{k=1}^{n} \mathbf{Var}[X_k]$$

By part (1),

$$\mathbf{Var}[X] = \sum_{k=1}^{n} \frac{1-p_k}{p_k^2}$$

If we let $j = n-k+1$, then $p_j = \frac{j}{n}$.

$$\begin{aligned}
&= \sum_{j=1}^{n} \frac{1-p_j}{p_j^2} \\
&= \sum_{j=1}^{n} \frac{1-\frac{j}{n}}{\frac{j^2}{n^2}} \\
&= \sum_{j=1}^{n} \frac{n^2-jn}{j^2} \\
&= n^2 \sum_{j=1}^{n} \frac{1}{j^2} - n \sum_{j=1}^{n} \frac{1}{j}
\end{aligned}$$

3. It turns out that the series $\sum_{i=1}^{\infty} \frac{1}{i^2}$ converges to a constant value $C = \frac{\pi^2}{6} \approx 1.645$. Deduce that $Var[X] \leq Cn^2$. Hence deduce the smallest value of β for which you can say that the probability we need to buy more than $\mu + \beta n$ boxes is less than $\frac{1}{100}$.

–Solution–

$$\begin{aligned}\mathbf{Var}[X] &= n^2 \sum_{j=1}^{n} \frac{1}{j^2} - n \sum_{j=1}^{n} \frac{1}{j} \\ &\leq n^2 \sum_{j=1}^{n} \frac{1}{j^2} \\ &\leq n^2 \sum_{j=1}^{\infty} \frac{1}{j^2} \\ &= n^2 C\end{aligned}$$

$$\begin{aligned}\mathbf{Pr}[X \geq \mu + \beta n] &= \mathbf{Pr}[(X - \mu) \geq \beta n] \\ &\leq \mathbf{Pr}[|X - \mu| \geq \beta n]\end{aligned}$$

Using Chebyshev's inequality,

$$\begin{aligned}&\leq \frac{\mathbf{Var}[X]}{\beta^2 n^2} \\ &\leq \frac{Cn^2}{\beta^2 n^2} \\ &= \frac{C}{\beta^2}\end{aligned}$$

This probability is smaller than $\frac{1}{100}$ when $\beta^2 \leq 100C$ — when $\beta \leq 10\sqrt{C} \approx 12.8$.

3 Exponential Distribution

1. Let r.v. X have exponential distribution with parameter λ. Show that, for any positive s, t we have $Pr[X > s + t | X > t] = Pr[X > s]$.
   NOTE: This is the memoryless property of the exponential distribution.

–Solution–

$$\mathbf{Pr}\left[X > s+t | X > t\right] = \frac{\mathbf{Pr}\left[X > s+t, X > t\right]}{\mathbf{Pr}\left[X > t\right]}$$

The event $X > s+t$ is a subset of $X > t$, therefore

$$\begin{aligned} &= \frac{\mathbf{Pr}\left[X > s+t\right]}{\mathbf{Pr}\left[X > t\right]} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} \\ &= e^{-\lambda s} \\ &= \mathbf{Pr}\left[X > s\right] \end{aligned}$$

2. Let r.v.s X_1, X_2 be independent and exponentially distributed with parameters λ_1, λ_2. Show that the r.v. $Y = min\{X_1, X_2\}$ is exponentially distributed with parameter $\lambda_1 + \lambda_2$.
Hint: work with the CDF (cumulative distribution function, $P[X \leq x]$).

–Solution–

$$\begin{aligned} \mathbf{Pr}\left[Y > t\right] &= \mathbf{Pr}\left[\min(X_1, X_2) > t\right] \\ &= \mathbf{Pr}\left[X_1 > t, X_2 > t\right] \\ &= \mathbf{Pr}\left[X_1 > t\right] \mathbf{Pr}\left[X_2 > t\right] \\ &= e^{-\lambda_1 t} e^{-\lambda_2 t} \\ &= e^{-(\lambda_1 + \lambda_2) t} \end{aligned}$$

This shows that Y has exponential distribution with parameter $\lambda_1 + \lambda_2$.

3. You have a digital camera that requires two batteries to operate. You purchase n batteries, labeled 1,2,...,n, each of which has a lifetime that is exponentially distributed with parameter λ and is independent of all the other batteries. Initially you install batteries 1 and 2. Each time a battery fails, you replace it with the lowest-numbered unused battery. At the end of this process you will be left with just one working battery. What is the expected total time until the end of the process? Justify your answer.

–Solution–

With two batteries in operation, the time until first failure is distributed as an exponential random variable with parameter 2λ, as in part (2). After a battery is replaced, because of the memoryless property of the exponential distribution, the time until next failure is

again distributed as an exponential random variable with parameter 2λ, as in part (1). Therefore, the total time until there is one battery left is $X = X_1 + X_2 + X_3 + \ldots + X_{n-1}$, and by linearity of expectation, $\mathbb{E}[X] = \frac{n-1}{2\lambda}$.

4 Gaussian

Let S be the number of Heads after flipping n coins, with $P[H] = p$; i.e., $S \approx Binomial(n, p)$. We have shown that $E[S] = np$ and $Var(S) = np(1-p)$. It turns out that, for large n, the binomial distribution is very closely approximated by the Gaussian distribution with the same mean and variance (this is a special case of the Central Limit Theorem). Use this information (and a suitable table or calculator for the cdf function of the Gaussian distribution) to answer the following:

1. Consider a coin with $P[H] = 9/10$. What is the (approximate) probability that, in 1000 independent flips of this coin, there will be at least 120 tails?

–Solution–

$\mathbf{E}[S] = np = 900$. $\mathbf{Var}[S] = np(1-p) = 90$. Using the above state special case of the central limit theorem, we are interested in $\mathbf{Pr}[S \leq 880]$, approximating S as a Gaussian. The z-score is $\frac{880-900}{\sqrt{90}} = -2.11$ Using a z-score table, $\Phi(-2.11) = 1 - \Phi(2.11) = 1 - 0.9826 = 0.0174$.

2. Find a value k such that, when you flip a fair coin 10,000 times, the probability of getting at least k heads is approximately 0.10.

–Solution–

$p = 0.5$, so if S represents the total number of heads, $\mathbf{E}[S] = 5000$, $\mathbf{Var}[S] = 2500$. The z-score for getting at least k heads is $\frac{k-5000}{50}$. From the z-score table, to have probability 0.10 of having at least k heads requires a z-score of roughly 1.28.
Therefore, $k \approx 5064$.

5 Random Variables in Galois Field

Let the random variables X and Y be distributed independently and uniformly at random in the set $\{0, 1, \ldots, p-1\}$, where $p > 2$ is a prime.

1. What is the expectation $\mathbf{E}[X]$?

–Solution–

X is a uniformly distribute random variable between 0 and $p-1$.

$$\mathbf{E}[X] = \frac{(p-1)-0}{2} = \frac{p-1}{2}$$

2. Let $S = (X+Y) \bmod p$ and $T = XY \bmod p$. What are the distributions of S and T?

–Solution–

For every a, and for every value of X we have a single value of Y, specifically $Y = a - X \pmod p$ such that $X + Y = a \pmod p$. This means that

$$\begin{aligned}
\mathbf{Pr}[S = a] &= \mathbf{Pr}[X + Y = a \pmod p] \\
&= \sum_{k=0}^{p-1} \mathbf{Pr}[X = k, Y = a - k \pmod p] \\
&= \sum_{k=0}^{p-1} \mathbf{Pr}[X = k]\,\mathbf{Pr}[Y = a - k \pmod p] \\
&= p \cdot \frac{1}{p} \cdot \frac{1}{p} \\
&= \frac{1}{p}
\end{aligned}$$

For nonzero a, and for every nonzero value of X we have a single value of Y, specifically $Y = aX^{-1} \pmod p$ such that $XY = a \pmod p$. This means that

$$\begin{aligned}
\mathbf{Pr}[T = a] &= \mathbf{Pr}[XY = a \pmod p] \\
&= \sum_{k=1}^{p-1} \mathbf{Pr}\left[X = k, Y = ak^{-1} \pmod p\right] \\
&= \sum_{k=1}^{p-1} \mathbf{Pr}[X = k]\,\mathbf{Pr}\left[Y = ak^{-1} \pmod p\right] \\
&= (p-1) \cdot \frac{1}{p} \cdot \frac{1}{p} \\
&= \frac{p-1}{p^2}
\end{aligned}$$

To make $XY = 0 \pmod{p}$, either $X = 0$ or $X \neq 0$ and $Y = 0$. Therefore

$$\begin{aligned}
\mathbf{Pr}[T = 0] &= \mathbf{Pr}[XY = 0 \pmod{p}] \\
&= \mathbf{Pr}[X = 0] + \mathbf{Pr}[X \neq 0, Y = 0] \\
&= \mathbf{Pr}[X = 0] + \mathbf{Pr}[X \neq 0]\,\mathbf{Pr}[Y = 0] \\
&= \frac{1}{p} + \frac{p-1}{p} \cdot \frac{1}{p} \\
&= \frac{2p-1}{p^2}
\end{aligned}$$

3. What are the expectations $\mathbf{E}[S]$ and $\mathbf{E}[T]$?

–Solution–

S is uniformly distributed over $0, \ldots, p-1$, making $\mathbf{E}[S] = \frac{p-1}{2}$. T is slightly more complicated:

$$\begin{aligned}
\mathbf{E}[T] &= \sum_{k=0}^{p-1} k\mathbf{Pr}[S = k] \\
&= 0 \cdot \mathbf{Pr}[S = 0] + \sum_{k=1}^{p-1} k\mathbf{Pr}[S = k] \\
&= \mathbf{Pr}[S = k] \sum_{k=1}^{p-1} k \\
&= \frac{p-1}{p^2} \cdot \frac{p(p-1)}{2} \\
&= \frac{(p-1)^2}{2p}
\end{aligned}$$

4. By linearity of expectation, we might expect that $\mathbf{E}[S] = (\mathbf{E}[X] + \mathbf{E}[Y]) \bmod p$. Explain why this does not hold in the present context; i.e. why does the value for $\mathbf{E}[S]$ obtained in part (c) not contradict linearity of expectation?

–Solution–

We have only shown that expectation is linear with respect to regular addition over the real numbers. We proved nothing about linearity with respect to addition modulu p. There is actually no reason to expect that $\mathbf{E}[X + Y \pmod{p}]$ will be the same as $\mathbf{E}[X] + \mathbf{E}[Y] \pmod{p}$.

–Solution–

The random variable $X : \Omega \to \mathbb{R}$, but confined to the interval $[0, p-1]$. The Galois Field defines a special addition operation, so the random variable $S : \Omega \to \mathbb{R}$ is again a random variable, so the output is constrained to $\mathbb{R}$. The addition operation under the Galois field is not a linear operation on the Reals.

5. Since X and Y are independent, we might expect that $\mathbf{E}[T] = \mathbf{E}[X]\mathbf{E}[Y] \bmod p$. Does this hold in this case? Explain why/why not.

–Solution–

Similarly, we only proved the result with respect to regular multiplication over the real numbers. We proved nothing about multiplication modulu p. There is no reason to expect that $\mathbf{E}[XY \pmod p]$ will be the same as $\mathbf{E}[X]\,\mathbf{E}[Y] \pmod p$.

6 Infinity

For each of the following sets, indicate whether it is finite, countably infinite, or uncountable. Give a one sentence justification for your choice:

1. Z (the set of all integers)

–Solution–

Countably infinite. Consider the function $f : \mathbb{N} \to \mathbb{Z}$ defined by $f(n) = (-1)^n \lceil \frac{n}{2} \rceil$, it is clearly a bijection between the naturals and the integers – as it enumerates them 0, -1, 1, -2, 2, etc.

2. Q (the set of all rational numbers)

–Solution–

Countably infinite. $|\mathbb{N}| \leq |\mathbb{Q}|$ by the identity function. $|\mathbb{Q}| \leq |\mathbb{Z} \times \mathbb{Z}|$ using $f(\frac{a}{b}) = (a, b)$. We've seen in class that $|\mathbb{Z} \times \mathbb{Z}|$ is countably infinite, so by Cantor-Bernestein the rationals are too.

3. R (the set of all real numbers)

–Solution–

Uncountable. Proved in class using Cantor's diagnolization argument.

4. C (the set of all complex numbers)

–Solution–

Uncountable. Using the identity function, $|\mathbb{R}| \leq |\mathbb{C}|$. Therefore, if C was countable, $\mathbb{R}$ would have to be countable.

5. $\{0,1\}^*$ (the set of all finite-length binary strings)

–Solution–

Countably infinite. We can just enumerate, in lexicographic ordering, all strings of length 0 (there's just one), then all strings of length 1 (there are two of them), then all strings of size 2, and so on. This defines a bijection between the naturals an all finite-length binary strings.

6. $\{0,1,2\}^*$ (the set of all finite-length ternary strings)

–Solution–

Countably infinite. Using the same argument as for all finite-length binary strings.

7. $\mathbb{Z}^3 = \{(a,b,c) : a,b,c \in Z\}$ (the set of triples of integers)

–Solution–

Countably infinite. We have seen in class that $\mathbb{Z}^2$ is countably infinite, we therefore have a bijection g between $\mathbb{Z}^2$ and $\mathbb{Z}$. Thus, the function $f((a,b,c)) = (g(a,b),c)$ is a bijection between $\mathbb{Z}^3$ and $\mathbb{Z}^2$, which is countably infinte.

8. $S = \{p(x) : p(x) = ax^2 + bx + c$, where $a,b,c \in Z\}$ (the set of all polynomials of degree at most 2 with integer coefficients)

–Solution–

Countably infinite. There is a simple bijection between S and $\mathbb{Z}^3$. Simply match the coefficients (a,b,c) to a tuple in $\mathbb{Z}^3$. Therefore, $|S| = |\mathbb{Z}^3|$.

9. $T = \{p(x) : p(x) = a_n x^n + ... + a_1 x + a_0$,where $n \in N$ and $a_0, a_1, \ldots, a_n \in Z\}$(the set of all polynomials with integer coefficients, of any degree)

–Solution–

Countably infinite. We show that $|T| \leq |\{0,1,2\}^*| = |\mathbb{N}|$ by encoding each polynomial as a ternary string; each coefficient a_k is encoded as a 0, a 2 if the a_k is negative, and then $|a_k|$ times 1. For example, $3x - 5$ is encoded as 01110211111. This is clearly a 1-1 encoding. $|T| \leq |\mathbb{N}|$ by the identity function. Applying Cantor-Bernstein gives us $|T| = |\mathbb{N}|$.

10. The set of all integer-valued random variables on a finite sample space.

–Solution–

Let $\Omega = \{\omega_1, \ldots, \omega_n\}$ be some finite set. An integer-valued random variable is function that maps an integer to every $\omega \in \Omega$. The function $f(X) = (X(\omega_1), \ldots, X(\omega_n))$ is a bijection from the integer-valued random variables to $\mathbb{Z}^n$, the set of all n-tuples of integers, which can be shown to be countably infinite by applying the solution of part (g) inductively.

7 Extra Credit

Suppose we have n samples $x_1 \ldots x_n$ drawn from some distribution. We wish to estimate the mean and variance of this distribution. The obvious choice for an estimate of the mean m is $\frac{x_1 + \cdots + x_n}{n}$. Convince yourself that $E[m] = E[x_i]$.

Now the estimate for the variance is

$$v = \frac{(x_1 - m)^2 + \cdots (x_n - m)^2}{n - 1}$$

Prove that $E[v] = E[x_i^2] - E[x_i]^2$. i.e. the expected value of v is the variance of the distribution we are sampling from.

–Solution–

$$\begin{aligned} v &= \frac{1}{n-1} \sum_{k=1}^{n} (x_k - m)^2 \\ &= \frac{1}{n-1} \sum_{k=1}^{n} \left[x_k^2 - 2mx_k + m \right] \\ &= \frac{1}{n-1} \left[\sum_{k=1}^{n} x_k^2 - 2m \sum_{k=1}^{n} x_k + \sum_{k=1}^{n} m^2 \right] \end{aligned}$$

we plug in the definition of m

$$= \frac{1}{n-1}\left[\sum_{k=1}^{n} x_k^2 - 2\left(\frac{\sum_{k=1}^{n} x_k}{n}\right)\sum_{k=1}^{n} x_k + n\left(\frac{\sum_{k=1}^{n} x_k}{n}\right)^2\right]$$
$$= \frac{1}{n-1}\left[\sum_{k=1}^{n} x_k^2 - \frac{1}{n}\left(\sum_{k=1}^{n} x_k\right)\left(\sum_{k=1}^{n} x_k\right)\right]$$

now take the expectation

$$\mathbf{E}[v] = \frac{1}{n-1}\left[\sum_{k=1}^{n} \mathbf{E}\left[x_k^2\right] - \frac{1}{n}\mathbf{E}\left[\left(\sum_{k=1}^{n} x_k\right)\left(\sum_{k=1}^{n} x_k\right)\right]\right]$$

Use linearity of expectation

$$= \frac{1}{n-1}\left[\sum_{k=1}^{n} \mathbf{E}\left[x_k^2\right] - \frac{1}{n}\sum_{1\le k,l\le n} \mathbf{E}[x_k x_l]\right]$$

Now gather all the terms together. All the samples are taken from the same distribution, but are independent, so $\mathbf{E}[x_k x_l] = \mathbf{E}\left[x_k^2\right] = \mathbf{E}\left[x_i^2\right]$ if $k = l$, or $\mathbf{E}[x_k x_l] = \mathbf{E}[x_k]\,\mathbf{E}[x_l] = \mathbf{E}[x_i]\,\mathbf{E}[x_i] = \mathbf{E}[x_i]^2$ if $k \neq l$

$$, = \frac{1}{n-1}\left[n\mathbf{E}\left[x_i^2\right] - \frac{1}{n}\left[n\mathbf{E}\left[x_i^2\right] + n(n-1)\mathbf{E}[x_i]^2\right]\right]$$
$$= \frac{1}{n-1}\left[(n-1)\mathbf{E}\left[x_i^2\right] - (n-1)\mathbf{E}[x_i]^2\right]$$
$$= \mathbf{E}\left[x_i^2\right] - \mathbf{E}[x_i]^2$$