Supplemental Document

In this supplemental document we prove the various properties and theorems referenced earlier (particularly those in Table 1).

Property 1. *If $F \cap H \subseteq G$ then $\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)] = \theta$.*

Proof.

$$\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)] \overset{(a)}{=} \mathbf{E}_g\left[\frac{f(X)}{g(X)}h(X)\right] = \int_G g(x)\frac{f(x)}{g(x)}h(x)\,\mathrm{d}x \overset{(b)}{=} \int_{F\cap H} f(x)h(x)\,\mathrm{d}x = \mathbf{E}_f[h(X)] = \theta,$$

where **(a)** holds because $\mathrm{IS}(\mathbf{X}_n)$ is the mean of n independent and identically distributed random variables, and **(b)** holds because $\forall x \in G \setminus (F \cap H)$, $f(x) = 0$. ■

We now provide a proof of Theorem 1, which states that if $C = G$, then $\mathrm{US}(\mathbf{X}_n) = \mathrm{IS}(\mathbf{X}_n)$.

Proof. In this setting, $c = \int_G g(x)\,\mathrm{d}x = 1$ and since every X_i must be within C, $k(\mathbf{X}_n) = n$. So,

$$\mathrm{US}(\mathbf{X}_n) = \frac{c}{k(\mathbf{X}_n)}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}h(X_i) = \frac{1}{n}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}h(X_i).$$ ■

We now provide a proof of Theorem 2, which states that if we replace c with an empirical estimate, $\hat{c}(\mathbf{X}_n) := n^{-1}k(\mathbf{X}_n)$, then $\mathrm{US}(\mathbf{X}_n) = \mathrm{IS}(\mathbf{X}_n)$.

Proof. Using the empirical estimate, $\hat{c}(\mathbf{X}_n)$, in place of c within US we have:

$$\begin{aligned}\mathrm{US}(\mathbf{X}_n) &= \frac{\hat{c}(\mathbf{X}_n)}{k(\mathbf{X}_n)}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}h(X_i)\\ &= \frac{k(\mathbf{X}_n)}{nk(\mathbf{X}_n)}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}h(X_i)\\ &= \frac{1}{n}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}h(X_i)\\ &= \mathrm{IS}(\mathbf{X}_n).\end{aligned}$$ ■

Theorem 3. *If $F \cap H \subseteq G$ and $\kappa \in \mathbb{N}_{>0}$, then*

$$\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n) = \kappa] = \theta.$$

Proof. Let $\Pr_g(X \in C)$ denote the probability that a sample, X, from the sampling distribution is in C.

$$\begin{aligned}&\mathbf{E}_g[\,\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n) = \kappa]\\ &= \mathbf{E}_g\left[\frac{c}{\kappa}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}h(X_i)\middle|k(\mathbf{X}_n) = \kappa\right]\\ &\overset{(a)}{=} \mathbf{E}_g\left[\frac{c}{\kappa}\sum_{i=1}^{\kappa}\frac{f(X_i)}{g(X_i)}h(X_i)\middle|\forall i \in \{1,\dots,\kappa\},\, X_i \in C\right]\\ &\overset{(b)}{=} \mathbf{E}_g\left[c\frac{f(X)}{g(X)}h(X)\middle|X \in C\right]\\ &\overset{(c)}{=} \int_C \frac{g(x)}{\Pr_g(X \in C)}c\frac{f(x)}{g(x)}h(x)\,\mathrm{d}x\\ &\overset{(d)}{=} \int_C \frac{g(x)}{c}c\frac{f(x)}{g(x)}h(x)\,\mathrm{d}x\\ &= \int_C f(x)h(x)\,\mathrm{d}x\\ &\overset{(e)}{=} \mathbf{E}_f[h(X)],\end{aligned}$$

where **(a)** holds because $f(X_i) = 0$ for all but κ of the terms in the summation, and so (by re-ordering the X_i so that these κ terms have indices $1, \dots, \kappa$) we need only sum to κ rather than n, **(b)** holds because the summation is over κ independent and identically distributed random variables, **(c)** holds by the definition of conditional expectations, **(d)** holds because $\Pr_g(X \in C) = c$, and **(e)** holds because $F \cap H \subseteq C$. ■

Theorem 4. *If $F \cap H \subseteq G$ then*

$$\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n) > 0] = \theta.$$

Proof.

$$\begin{aligned}&\mathbf{E}_g[\,\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n) > 0]\\ &= \sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n) = \kappa|k(\mathbf{X}_n) > 0)}{\Pr(k(\mathbf{X}_n) > 0)}\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n) = \kappa]\\ &\overset{(a)}{=} \sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n) = \kappa|k(\mathbf{X}_n) > 0)}{\Pr(k(\mathbf{X}_n) > 0)}\theta\\ &= \theta\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n) = \kappa|k(\mathbf{X}_n) > 0)}{\Pr(k(\mathbf{X}_n) > 0)}\\ &= \theta,\end{aligned}$$

where **(a)** holds because, by Theorem 3, $\mathbf{E}[\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n) = \kappa] = \theta$. ■

Theorem 5. *If $F \cap H \subseteq G$ and $\kappa \in \mathbb{N}_{>0}$, then*

$$\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n) = \kappa] - \theta = \left(\frac{\kappa}{cn} - 1\right)\theta. \quad (4)$$

Proof. Following roughly the same steps as used to prove

Theorem 3 we have that:

$$\begin{aligned}
&\mathbf{E}_g[\,\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n)=\kappa]\\
=&\mathbf{E}_g\left[\frac{1}{n}\sum_{i=1}^{n}\frac{f(X_i)}{g(X_i)}h(X_i)\middle|k(\mathbf{X}_n)=\kappa\right]\\
=&\mathbf{E}_g\left[\frac{1}{n}\sum_{i=1}^{\kappa}\frac{f(X_i)}{g(X_i)}h(X_i)\middle|\forall i\in\{1,\dots,\kappa\},X_i\in C\right]\\
=&\mathbf{E}_g\left[\frac{\kappa}{n}\frac{f(X_1)}{g(X_1)}h(X_1)\middle|X_1\in C\right]\\
=&\int_C\frac{g(x)}{c}\frac{\kappa}{n}\frac{f(x)}{g(x)}h(x)\,\mathrm{d}x\\
=&\frac{\kappa}{cn}\mathbf{E}_f[h(X)]\\
=&\frac{\kappa}{cn}\theta,
\end{aligned}$$

and so (4) follows. ■

Theorem 6. *If $F\cap H\subseteq G$ then*

$$\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n)>0]=\frac{1}{1-(1-c)^n}\theta.$$

Proof. Recall from Property 1 that $\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)]=\theta$. By marginalizing over whether or not $k(\mathbf{X}_n)>0$, we also have that:

$$\begin{aligned}
\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)]=&\Pr(k(\mathbf{X}_n)>0)\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n)>0]\\
&+\Pr(k(\mathbf{X}_n)=0)\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n)=0].
\end{aligned}$$

So,

$$\begin{aligned}
&\mathbf{E}_g[\,\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n)>0]\\
&=\frac{\theta-\Pr(k(\mathbf{X}_n)=0)\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n)=0]}{\Pr(k(\mathbf{X}_n)>0)}\\
&\overset{\text{(a)}}{=}\frac{\theta}{1-(1-c)^n},
\end{aligned}$$

where **(a)** holds because $\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n)=0]=0$ and $\Pr(k(\mathbf{X}_n)>0)=1-\Pr(k(\mathbf{X}_n)=0)=1-(1-c)^n$. ■

Theorem 7. *If $F\cap H\subseteq G$, then*

$$\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)]=(1-(1-c)^n)\theta.$$

Proof.

$$\begin{aligned}
&\mathbf{E}_g[\,\mathrm{US}(\mathbf{X}_n)]\\
&=\underbrace{\Pr(k(\mathbf{X}_n)>0)}_{=1-(1-c)^n}\underbrace{\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n)>0]}_{=\theta,\text{ by Theorem 4}}\\
&\quad+\Pr(k(\mathbf{X}_n)=0)\underbrace{\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n)=0]}_{=0}\\
&=(1-(1-c)^n)\theta.
\end{aligned}$$

■

Before continuing, recall the following property (which we prove for completeness):

Property 2. *Let $X_1,\dots,X_n$ be n independent and identically distributed random variables, each with finite mean and variance. Then,*

$$\mathbf{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)^2\right]=\frac{1}{n}\,\mathrm{Var}\,(X_1)+\mathbf{E}\,[X_1]^2\,.$$

Proof. Recall that

$$\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\mathbf{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)^2\right]-\mathbf{E}\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]^2.$$

So, by rearranging terms:

$$\mathbf{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)^2\right]=\frac{1}{n^2}\,\mathrm{Var}\left(\sum_{i=1}^{n}X_i\right)+\frac{1}{n^2}\mathbf{E}\left[\sum_{i=1}^{n}X_i\right]^2.$$

Since the X_i are independent and identically distributed, we therefore have that:

$$\begin{aligned}
\mathbf{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)^2\right]=&\frac{1}{n^2}n\,\mathrm{Var}\,(X_1)+\frac{1}{n^2}n^2\mathbf{E}\,[X_1]^2\\
=&\frac{1}{n}\,\mathrm{Var}\,(X_1)+\mathbf{E}\,[X_1]^2\,.
\end{aligned}$$

■

Theorem 8. *If $F\cap H\subseteq G$ then*

$$\mathrm{Var}_g(\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n>0))=c^2v\mathbf{E}_{B(n,c)}\left[\frac{1}{\kappa}\middle|\kappa>0\right].$$

Proof.

$$\begin{aligned}
&\mathrm{Var}_g(\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n)>0)\\
&=\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)^2|k(\mathbf{X}_n)>0]-\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n)>0]^2\\
&=\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)^2|k(\mathbf{X}_n)>0]-\theta^2\\
&=\left(\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n)=\kappa)}{\Pr(k(\mathbf{X}_n)>0)}\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)^2|k(\mathbf{X}_n)=\kappa]\right)-\theta^2.
\end{aligned}\tag{5}$$

We will write $\mathbf{y}$ to denote a vector in $\mathbb{R}^n$, the elements of which are $y_1,\dots,y_n\in\mathbb{R}$. We also write $\mathbf{y}_{i:j}$ to denote the i^{th} through j^{th} entries of $\mathbf{y}$, i.e., $\mathbf{y}_{i:j}:=[y_i,y_{i+1},\dots,y_{j-1},y_j]$. Let $G^n_\kappa=\{\mathbf{y}\in G^n:k(\mathbf{y})=\kappa\}$ be the set of all possible tuples of n samples where exactly κ are in C. We also overload the definition of g by defining $g(\mathbf{y}):=\prod_{i=1}^{n}g(y_i)$. Using this notation, we have that (where $\dots$ are used to denote that a long line is split across multiple lines via scalar multiplication):

$$
\begin{aligned}
&\mathbf{E}_g[\,\mathrm{US}(\mathbf{X}_n)^2|k(\mathbf{X}_n)=\kappa]\\
&=\int_{G_\kappa^n}\frac{g(\mathbf{y})}{\Pr(k(\mathbf{X}_n)=\kappa)}\,\mathrm{US}(\mathbf{y})^2\,\mathrm{d}\mathbf{y}\\
&\overset{\text{(a)}}{=}\frac{\binom{n}{\kappa}}{\Pr(k(\mathbf{X}_n)=\kappa)}\int_{C^\kappa}\int_{(G\setminus C)^{n-\kappa}}g(\mathbf{y})\,\mathrm{US}(\mathbf{y})^2\,\mathrm{d}\mathbf{y}_{1:\kappa}\,\mathrm{d}\mathbf{y}_{\kappa+1:n}\\
&\overset{\text{(b)}}{=}\frac{\binom{n}{\kappa}}{\Pr(k(\mathbf{X}_n)=\kappa)}\int_{C^\kappa}\int_{(G\setminus C)^{n-\kappa}}g(\mathbf{y}_{1:\kappa})g(\mathbf{y}_{\kappa+1:n})\dots\\
&\quad\mathrm{US}(\mathbf{y}_{1:\kappa})^2\,\mathrm{d}\mathbf{y}_{1:\kappa}\,\mathrm{d}\mathbf{y}_{\kappa+1:n}\\
&=\frac{\binom{n}{\kappa}}{\binom{n}{\kappa}c^\kappa(1-c)^{n-\kappa}}\int_{C^\kappa}g(\mathbf{y}_{1:\kappa})\,\mathrm{US}(\mathbf{y}_{1:\kappa})^2\,\mathrm{d}\mathbf{y}_{1:\kappa}\dots\\
&\quad\underbrace{\int_{(G\setminus C)^{n-\kappa}}g(\mathbf{y}_{\kappa+1:n})\,\mathrm{d}\mathbf{y}_{\kappa+1:n}}_{=(1-c)^{n-\kappa}}\\
&=\frac{\binom{n}{\kappa}(1-c)^{n-k}}{\binom{n}{\kappa}c^\kappa(1-c)^{n-\kappa}}\int_{C^\kappa}g(\mathbf{y}_{1:\kappa})\left(\frac{c}{\kappa}\sum_{i=1}^{\kappa}\frac{f(y_i)}{g(y_i)}h(y_i)\right)^2\mathrm{d}\mathbf{y}_{1:\kappa}\\
&=\frac{c^2}{c^\kappa}\int_{C^\kappa}g(\mathbf{y}_{1:\kappa})\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}\frac{f(y_i)}{g(y_i)}h(y_i)\right)^2\mathrm{d}\mathbf{y}_{1:\kappa}\\
&\overset{\text{(c)}}{=}c^2\int_{C^\kappa}\frac{g(\mathbf{y}_{1:\kappa})}{\Pr(k(\mathbf{X}_\kappa)=\kappa)}\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}\frac{f(y_i)}{g(y_i)}h(y_i)\right)^2\mathrm{d}\mathbf{y}_{1:\kappa}\\
&=c^2\mathbf{E}_g\left[\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}\frac{f(X_i)}{g(X_i)}h(X_i)\right)^2\middle|\mathbf{X}_\kappa\in C^\kappa\right]\\
&\overset{\text{(d)}}{=}c^2\left(\frac{1}{\kappa}v+\mathbf{E}\left[\frac{f(X)}{g(X)}h(X)\middle|X\sim g,X\in C\right]^2\right)\\
&=c^2\left(\frac{1}{\kappa}v+\left(\int_C\frac{g(x)}{c}\frac{f(x)}{g(x)}h(x)\,\mathrm{d}x\right)^2\right)=\frac{c^2}{\kappa}v+\theta^2, \qquad (6)
\end{aligned}
$$

where **(a)** comes from **1)** the fact that there are $\binom{n}{\kappa}$ ways of ordering n elements such that κ are in C and $n-\kappa$ are in $G\setminus C$, and **2)** the fact that US does not depend on the order of its inputs, **(b)** comes from **1)** the property that $\mathrm{US}(\mathbf{y})$ does not change if additional samples are appended to $\mathbf{y}$ that are not in C and **2)** the fact that $g(\mathbf{y})$ can be decomposed into $g(\mathbf{y}_{1:\kappa})g(\mathbf{y}_{\kappa+1:n})$ since it represents the joint probability density function for n independent and identically distributed random variables, **(c)** comes from the fact that $\Pr(k(\mathbf{X}_\kappa)=\kappa)=c^\kappa$, and **(d)** comes from Property 2.

Combining (5) with (6) we have that

$$
\begin{aligned}
&\mathrm{Var}_g(\mathrm{US}(\mathbf{X}_n)|k(\mathbf{X}_n)>0)\\
&=\left(\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n)=\kappa)}{\Pr(k(\mathbf{X}_n)>0)}\left(\frac{c^2}{\kappa}v+\theta^2\right)\right)-\theta^2\\
&=c^2v\left(\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n)=\kappa)}{\Pr(k(\mathbf{X}_n)>0)}\frac{1}{\kappa}\right)\\
&\quad+\theta^2\underbrace{\left(\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n)=\kappa)}{\Pr(k(\mathbf{X}_n)>0)}\right)}_{=1}-\theta^2\\
&=c^2v\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n)=\kappa)}{\Pr(k(\mathbf{X}_n)>0)}\frac{1}{\kappa}\\
&=c^2v\mathbf{E}_{B(n,c)}\left[\frac{1}{\kappa}\middle|\kappa>0\right].
\end{aligned}
$$

■

Theorem 9. *If $F\cap H\subseteq G$ then*

$$
\mathrm{Var}_g(\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n>0))=v\frac{c}{n\rho}+\theta^2\frac{c\rho(n-1)+\rho-cn}{cn\rho^2}.
$$

Proof. At a high level, this proof is similar to the proof of Theorem 8, but uses the property that $\mathrm{IS}(\mathbf{X}_n)=\frac{k(\mathbf{X}_n)}{cn}\,\mathrm{US}(\mathbf{X}_n)$.

$$
\begin{aligned}
&\mathrm{Var}_g(\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n)>0)\\
&=\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)^2|k(\mathbf{X}_n)>0]-\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n)>0]^2\\
&\overset{\text{(a)}}{=}\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)^2|k(\mathbf{X}_n)>0]-\left(\frac{\theta}{1-(1-c)^n}\right)^2\\
&=\left(\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n)=\kappa)}{\Pr(k(\mathbf{X}_n)>0)}\mathbf{E}_g[\mathrm{IS}(\mathbf{X}_n)^2|k(\mathbf{X}_n)=\kappa]\right)\\
&\quad-\left(\frac{\theta}{1-(1-c)^n}\right)^2, \qquad (7)
\end{aligned}
$$

where **(a)** comes from Theorem 6.

Also,

$$
\begin{aligned}
&\mathbf{E}_g[\,\mathrm{IS}(\mathbf{X}_n)^2|k(\mathbf{X}_n)=\kappa]\\
&\overset{\text{(a)}}{=}\mathbf{E}_g\left[\left(\frac{k(\mathbf{X}_n)}{cn}\,\mathrm{US}(\mathbf{X}_n)\right)^2\middle|k(\mathbf{X}_n)=\kappa\right]\\
&=\frac{\kappa^2}{c^2n^2}\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)^2|k(\mathbf{X}_n)=\kappa]\overset{\text{(b)}}{=}\frac{\kappa^2}{c^2n^2}\left(\frac{c^2}{\kappa}v+\theta^2\right), \qquad (8)
\end{aligned}
$$

where **(a)** holds because $\mathrm{IS}(\mathbf{X}_n)=\frac{k(\mathbf{X}_n)}{cn}\,\mathrm{US}(\mathbf{X}_n)$ and **(b)** follows from (6). Using the shorthand, $\rho:=\Pr(k(\mathbf{X}_n)>0)=1-(1-c)^n$ and by combining (7) with (8) we have that:

$$
\begin{aligned}
&\mathrm{Var}_g(\mathrm{IS}(\mathbf{X}_n)|k(\mathbf{X}_n)>0)\\
&=\left(\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n)=\kappa)}{\Pr(k(\mathbf{X}_n)>0)}\frac{\kappa^2}{c^2n^2}\left(\frac{c^2}{\kappa}v+\theta^2\right)\right)\\
&\quad-\left(\frac{\theta}{1-(1-c)^n}\right)^2\\
&=\frac{v}{n^2\rho}\underbrace{\left(\sum_{\kappa=1}^{n}\Pr(k(\mathbf{X}_n)=\kappa)\kappa\right)}_{=\mathbf{E}_{B(n,c)}[\kappa]=nc}\\
&\quad+\frac{\theta^2}{c^2n^2\rho}\underbrace{\left(\sum_{\kappa=1}^{n}\Pr(k(\mathbf{X}_n)=\kappa)\kappa^2\right)}_{=\mathbf{E}_{B(n,c)}[\kappa^2]=nc((n-1)c+1)}-\left(\frac{\theta}{\rho}\right)^2\\
&=v\frac{c}{n\rho}+\frac{\theta^2((n-1)c+1)}{cn\rho}-\frac{\theta^2}{\rho^2}\\
&=v\frac{c}{n\rho}+\theta^2\frac{c\rho(n-1)+\rho-cn}{cn\rho^2}.
\end{aligned}
$$

■

Theorem 10. *If $F\cap H\subseteq G$ then*

$$\mathrm{Var}_g(\mathrm{US}(\mathbf{X}_n))=\rho c^2 v\mathbf{E}_{B(n,c)}\left[\frac{1}{\kappa}\middle|\kappa>0\right]+\theta^2\rho(1-\rho).$$

Proof.

$$
\begin{aligned}
&\mathrm{Var}_g(\mathrm{US}(\mathbf{X}_n))=\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)^2]-\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)]^2\\
&\overset{(a)}{=}\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)^2]-\rho^2\theta^2\\
&=\left(\sum_{\kappa=0}^{n}\Pr(k(\mathbf{X}_n)=\kappa)\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)^2|k(\mathbf{X}_n)=\kappa]\right)\\
&\quad-\rho^2\theta^2\\
&=\Pr(k(\mathbf{X}_n)=0)\underbrace{\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)^2|k(\mathbf{X}_n)=0]}_{=0}\\
&\quad+\left(\sum_{\kappa=1}^{n}\Pr(k(\mathbf{X}_n)=\kappa)\mathbf{E}_g[\mathrm{US}(\mathbf{X}_n)^2|k(\mathbf{X}_n)=\kappa]\right)\\
&\quad-\rho^2\theta^2\\
&\overset{(b)}{=}\rho\left(\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n)=\kappa)}{\rho}\left(\frac{c^2}{\kappa}v+\theta^2\right)\right)-\rho^2\theta^2\\
&=\rho c^2 v\left(\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n)=\kappa)}{\rho}\frac{1}{\kappa}\right)\\
&\quad+\rho\theta^2\underbrace{\left(\sum_{\kappa=1}^{n}\frac{\Pr(k(\mathbf{X}_n)=\kappa)}{\rho}\right)}_{=1}-\rho^2\theta^2\\
&=\rho c^2 v\mathbf{E}_{B(n,c)}\left[\frac{1}{\kappa}\middle|\kappa>0\right]+\theta^2\rho(1-\rho),
\end{aligned}
$$

where **(a)** comes from Theorem 7, **(b)** comes from (6) and from multiplying one term by $\rho/\rho=1$.

Theorem 11. *If $F\cap H\subseteq G$ then*

$$\mathrm{Var}_g(\mathrm{IS}(\mathbf{X}_n))=\frac{1}{n}\left(cv+\theta^2\left(\frac{1}{c}-1\right)\right).$$

Proof.

$$
\begin{aligned}
&\mathrm{Var}_g(\mathrm{IS}(\mathbf{X}_n))\overset{(a)}{=}\frac{1}{n}\mathrm{Var}_g(\mathrm{IS}(X))\\
&=\frac{1}{n}\left(\mathbf{E}_g[\mathrm{IS}(X)^2]-\mathbf{E}_g[\mathrm{IS}(X)]^2\right)\\
&\overset{(b)}{=}\frac{1}{n}\left(\mathbf{E}_g[\mathrm{IS}(X)^2]-\theta^2\right)\\
&=\frac{1}{n}\Bigg(\Pr(X\in C|X\sim g)\mathbf{E}_g[\mathrm{IS}(X)^2|X\in C]\\
&\quad+\Pr(X\notin C|X\sim g)\underbrace{\mathbf{E}_g[\mathrm{IS}(X)^2|X\notin C]}_{=0}-\theta^2\Bigg)\\
&=\frac{1}{n}\left(c\mathbf{E}_g[\mathrm{IS}(X)^2|X\in C]-\theta^2\right)\\
&\overset{(c)}{=}\frac{1}{n}\left(c\left(v+\frac{\theta^2}{c^2}\right)-\theta^2\right)\\
&=\frac{1}{n}\left(cv+\theta^2\left(\frac{1}{c}-1\right)\right),
\end{aligned}
$$

where **(a)** holds because $\mathrm{IS}(\mathbf{X}_n)$ is the sum of n independent and identically distributed random variables, **(b)** comes from Property 1, and **(c)** comes from applying (8) with $n=1$ and $\kappa=1$. ■

Property 3. *$c\rho(n-1)+\rho-cn\geq 0$,*

Proof. Recall that $\rho\coloneqq 1-(1-c)^n$, so we have that:

$$
\begin{aligned}
&c\rho(n-1)+\rho-cn=c(1-(1-c)^n)(n-1)+1-(1-c)^n-cn\\
&=(cn-c)(1-(1-c)^n)+1-(1-c)^n-cn\\
&=cn-cn(1-c)^n-c+c(1-c)^n+1-(1-c)^n-cn\\
&=(1-c)^n(-cn+c-1)-c+1. \qquad (9)
\end{aligned}
$$

We will show by induction that (9) is non-negative for all $n\geq 1$. First, notice that for the base case where $n=1$, (9) is equal to zero. For the inductive step we will show that (9) is non-negative for $n+1$ given that it is non-negative for n.

$$
\begin{aligned}
&(1-c)^{n+1}(-c(n+1)+c-1)-c+1\\
&\quad=(1-c)(1-c)^n(-cn+c-1)-(1-c)^{n+1}c\\
&\qquad+(-c+1)(1-c+c)\\
&\quad=(1-c)\underbrace{\left((1-c)^n(-cn+c-1)-c+1\right)}_{(a)}\\
&\qquad-(1-c)^{n+1}c+c(1-c),
\end{aligned}
$$

where **(a)** is positive by the inductive hypothesis, and so we need only show that $-(1-c)^{n+1}c+c(1-c)\geq 0$. Since

$$-(1-c)^{n+1}c+c(1-c)=c\Big((1-c)-(1-c)^{n+1}\Big),$$

and $1-c\geq(1-c)^{n+1}$ because $c\in(0,1]$, we conclude. ■