1. Låt $X_i \in N(3,\ 0.5)$ vara den mängd fotogen Muminfamiljen använder dag i.

   (a) På tre veckor, dvs $n = 21$ dagar, använder Muminfamiljen $Y = \sum_{i=1}^{n} X_i$ dl fotogen. Då har vi att
   $\mathsf{E}(Y) = \mathsf{E}(\sum_{i=1}^{21} X_i) = \sum_{i=1}^{21} \mathsf{E}(X_i) = \sum_{i=1}^{21} 3 = 21 \cdot 3 = 63$ dl medan
   $\mathsf{V}(Y) = \mathsf{V}(\sum_{i=1}^{21} X_i) = [\text{eftersom } X_i \text{ är oberoende}] = \sum_{i=1}^{21} \mathsf{V}(X_i) = \sum_{i=1}^{21} 0.5^2 = 21 \cdot 0.5^2\ \text{dl}^2$ och
   $\mathsf{D}(Y) = \sqrt{\mathsf{V}(Y)} = 0.5\sqrt{21}$ dl.
   Eftersom X_i är normalfördelade gäller det också att $Y \in N\left(63,\ 0.5\sqrt{21}\right)$.

   (b) Låt μ vara mängden fotogen i dunken. Muminpappan har konstaterat att $\mathsf{P}(Y \le \mu) = 0.95$. Eftersom $\mathsf{P}(Y \le \mu) = \mathsf{P}(\frac{Y-63}{0.5\sqrt{21}} \le \frac{\mu-63}{0.5\sqrt{21}}) = \Phi(\frac{\mu-63}{0.5\sqrt{21}}) = 0.95$ får vi, genom att utnyttja att $\Phi(\lambda_\alpha) = 1-\alpha$, att $\frac{\mu-63}{0.5\sqrt{21}} = \lambda_{1-0.95} = \lambda_{0.05}$ så att $\mu = 63 + \underbrace{\lambda_{0.05}}_{1.64} \cdot 0.5\sqrt{21} \approx 66.77$ dl.
   Dunken rymmer ca 6.7 liter.

   (c) Den verkliga fotogenåtgången är $Z = Y + 6 \in N\left(63+6,\ 0.5\sqrt{21}\right) = N\left(69,\ 0.5\sqrt{21}\right)$ och
   $\mathsf{P}(Z \le \mu) = \Phi(Z \le 66.77) = \Phi(\frac{66.77-69}{0.5\sqrt{12}}) = \Phi(-0.97) = 1-\Phi(0.97) = 1-0.8349 = 0.1651.$

2. Eftersom vi ska studera ett linjärt samband ansätter vi den linjära regressionsmodellen $y_i = \alpha + \beta x_i + \varepsilon_i$ där $\varepsilon_i \in N(0,\ \sigma)$ är oberoende för $i = 1, \ldots, n$ med $n = 10$. Om det finns ett linjärt samband kommer β att vara skilt från noll. Vi ska alltså testa hypoteserna $H_0\colon \beta = 0$ mot $H_1\colon \beta \neq 0$ på signifikansnivån, t.ex., $p = 0.05$.

   Vi har att $\beta^* = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2} = \frac{S_{xy}}{S_{xx}} = \frac{-1.046}{1.416} = -0.7387$ där $\beta^* \in N\left(\beta,\ \frac{\sigma}{\sqrt{S_{xx}}}\right)$ och
   $$\sigma^* = s = \sqrt{\frac{Q_0}{n-2}} = \left[\text{med } Q_0 = S_{yy} - \frac{S_{xy}^2}{S_{xx}} = 4.796 - \frac{(-1.046)^2}{1.416} = 4.0233\right] = \sqrt{\frac{4.0233}{10-2}} = 0.7092.$$

   Alt.1: Konfidensintervall. Eftersom $\beta^* \in N\left(\beta,\ \frac{\sigma}{\sqrt{S_{xx}}}\right)$, där σ har skattats med s, ges ett tvåsidigt 95 % konfidensintervall för β av $I_\beta = (\beta^* \pm t_{p/2}(n-2)\cdot\frac{s}{\sqrt{S_{xx}}}) = (-0.7387 \pm \underbrace{t_{0.025}(8)}_{2.31} \cdot \frac{0.7092}{\sqrt{1.416}}) = (-2.11,\ 0.64)$.
   Eftersom $H_0\colon \beta = 0$ ligger i intervallet kan H_0 inte förkastas. Det finns alltså inget signifikant linjärt samband mellan köldutbredning och storlek hos mårror.

   Alt.2: Teststorhet. Om $H_0\colon \beta = 0$ är sann gäller att $\beta^* \in N\left(0,\ \frac{\sigma}{\sqrt{S_{xx}}}\right)$ och därmed också att $\frac{\beta^*-0}{\sigma/\sqrt{S_{xx}}} \in N(0,\ 1)$ och $t = \frac{\beta^*-0}{s/\sqrt{S_{xx}}} \in t(n-2)$. Vi vill undersöka om β^* ligger misstänkt långt från 0, dvs om $|t|$ är större än det borde vara.
   Eftersom $|t| = \left|\frac{\beta^*-0}{s/\sqrt{S_{xx}}}\right| = \left|\frac{-0.7387-0}{0.7092/\sqrt{1.416}}\right| = |-1.2395| = 1.2395 \not> t_{0.025}(n-2) = t_{0.025}(8) = 2.31$ kan H_0 inte förkastas. Det finns alltså inget signifikant linjärt samband mellan köldutbredning och storlek hos mårror.

3. Vi har, enligt formelsamlingen, att när $X \in R(a,\ b) = R(-1,\ 1)$ så gäller att $E(X) = \dfrac{a+b}{2} = \dfrac{-1+1}{2} = 0$ och $\mathsf{V}(X) = \dfrac{(b-a)^2}{12} = \dfrac{(1-(-1))^2}{12} = \dfrac{1}{3}$ samt $f_X(x) = \dfrac{1}{b-a} = \dfrac{1}{1-(-1)} = \dfrac{1}{2}$ för $a \leq x \leq b$, dvs för $-1 \leq x \leq 1$.

   Enligt satsen om väntevärden av funktioner av stokastiska variabler har vi att

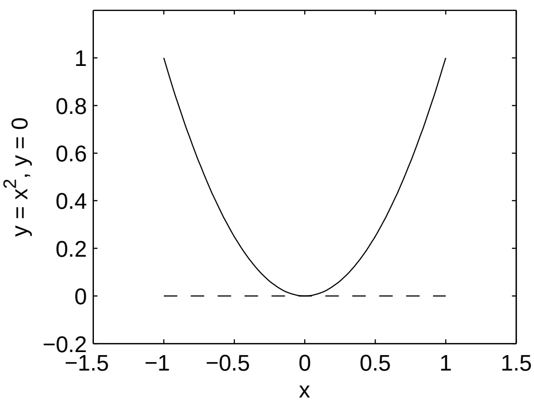
$$\mathsf{E}(Y) = \mathsf{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f_X(x)\,dx = \int_{-1}^{1} \frac{x^2}{2}\,dx = \left[\frac{x^3}{6}\right]_{-1}^{1} = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

   Detta kunde också beräknats med $\mathsf{V}(X) = \mathsf{E}(X^2) - \mathsf{E}^2(X) \Rightarrow \mathsf{E}(X^2) = \mathsf{V}(X) + \mathsf{E}^2(X) = \dfrac{1}{3} + 0^2 = \dfrac{1}{3}$.

   Eftersom $\mathsf{E}(Y^2) = \mathsf{E}(X^4) = \displaystyle\int_{-1}^{1} x^4 \cdot f_X(x)\,dx = \int_{-1}^{1} \frac{x^4}{2}\,dx = \left[\frac{x^5}{10}\right]_{-1}^{1} = \frac{1}{10} + \frac{1}{10} = \frac{1}{5}$ får vi att $\mathsf{V}(Y) = \mathsf{E}(Y^2) - \mathsf{E}^2(Y) = \dfrac{1}{5} - (\dfrac{1}{3})^2 = \dfrac{4}{45}$.

   Med Gaussapproximation får vi istället $\mathsf{E}(Y) = \mathsf{E}(X^2) \approx \mathsf{E}^2(X) = 0^2 = 0$ respektive $\mathsf{V}(Y) = \mathsf{V}(X^2) = \left[\text{eftersom } \dfrac{d}{dx}x^2 = 2x\right] = (2\mathsf{E}(X))^2 \cdot \mathsf{V}(X) = (2 \cdot 0)^2 \cdot \dfrac{1}{3} = 0.$

   Gaussapproximationen utnyttjar första ordningens Taylorutveckling av $y = g(x)$ kring punkten $x = \mathsf{E}(X)$, dvs $g(x) = x^2 \approx$
   $\approx g(\mathsf{E}(X)) + g'(\mathsf{E}(X)) \cdot \dfrac{(x - \mathsf{E}(X))^1}{1!} =$
   $= \mathsf{E}^2(X) + 2\mathsf{E}(X) \cdot (x - \mathsf{E}(X))$.
   Eftersom $\mathsf{E}(X) = 0$ approximeras funktionen $y = x^2$ (heldragen) med funktionen $y = 0$ (streckad). Det är en värdelös approximation eftersom x^2 är långt ifrån linjär i närheten av $x = \mathsf{E}(X) = 0$.



4. Låt X vara antalet hatifnattar på en $t\,\text{dm}^2$ stor gräsmatta. Det gäller, enligt uppgiften, att $X \in Po(\lambda t)$ samt att antalet hatifnattar på olika markbitar är oberoende. Vi har också nytta av att $\mathsf{E}(X) = \mathsf{V}(X) = \lambda t$.

   (a) Vi har att $t = 1\,\text{m}^2 = 100\,\text{dm}^2$ så att $X \in Po(\lambda t) = Po(1 \cdot 100) = Po(100)$.
   Eftersom $\lambda t = 100 > 15$ kan vi utnyttja normalapproximation och
   $X \underset{\sim}{\in} N(\mathsf{E}(X),\ \mathsf{D}(X)) = N\left(100,\ \sqrt{100}\right) = N(100,\ 10)$.
   Då blir $\mathsf{P}(X > 120) = 1 - \mathsf{P}(X \leq 120) = 1 - \mathsf{P}(\dfrac{X-100}{10} \leq \dfrac{120-100}{10}) \approx$
   $\approx 1 - \Phi(\dfrac{120-100}{10}) = 1 - \Phi(2) \approx 1 - 0.9772 = 0.0228.$

   (b) Vi har oberoende observationer $x_{\text{Sm}} = 4$, $x_{\text{Pv}} = 91$, $x_{\text{Pt}} = 51$ av de stokastiska variablerna $X_{\text{Sm}} \in Po(5\lambda)$, $X_{\text{Pv}} \in Po(100\lambda)$ respektive $X_{\text{Pt}} \in Po(50\lambda)$ och ska jämföra skattningarna
   $\lambda^*_{\text{Pv}} = \dfrac{1}{3}(\dfrac{X_{\text{Sm}}}{5} + \dfrac{X_{\text{Pv}}}{100} + \dfrac{X_{\text{Pt}}}{50})$, $\lambda^*_{\text{Sm}} = \dfrac{X_{\text{Sm}} + X_{\text{Pv}} + X_{\text{Pt}}}{5+100+50}$ och $\lambda^*_{\text{Pt}} = \dfrac{X_{\text{Pv}} + X_{\text{Pt}}}{100+50}$.
   Alla tre skattningarna är väntevärdesriktiga eftersom

$$\mathsf{E}(\lambda^*_{\text{Pv}}) = \mathsf{E}(\frac{1}{3}(\frac{X_{\text{Sm}}}{5} + \frac{X_{\text{Pv}}}{100} + \frac{X_{\text{Pt}}}{50})) = \frac{1}{3}(\frac{\mathsf{E}(X_{\text{Sm}})}{5} + \frac{\mathsf{E}(X_{\text{Pv}})}{100} + \frac{\mathsf{E}(X_{\text{Pt}})}{50}) = \frac{1}{3}(\frac{5\lambda}{5} + \frac{100\lambda}{100} + \frac{50\lambda}{50}) = \lambda,$$

$$\mathsf{E}(\lambda^*_{\text{Sm}}) = \mathsf{E}(\frac{X_{\text{Sm}} + X_{\text{Pv}} + X_{\text{Pt}}}{5+100+50}) = \frac{\mathsf{E}(X_{\text{Sm}}) + \mathsf{E}(X_{\text{Pv}}) + \mathsf{E}(X_{\text{Pt}})}{5+100+50} = \frac{5\lambda + 100\lambda + 50\lambda}{5+100+50} = \lambda,$$

$$\mathsf{E}(\lambda^*_{\text{Pt}}) = \mathsf{E}(\frac{X_{\text{Pv}} + X_{\text{Pt}}}{100+50}) = \frac{\mathsf{E}(X_{\text{Pv}}) + \mathsf{E}(X_{\text{Pt}})}{100+50} = \frac{100\lambda + 50\lambda}{100+50} = \lambda.$$

Deras varianser är

$$V(\lambda^*_{Pv}) = V(\frac{1}{3}(\frac{X_{Sm}}{5} + \frac{X_{Pv}}{100} + \frac{X_{Pt}}{50})) = [\text{ober.}] = \frac{1}{3^2}(\frac{V(X_{Sm})}{5^2} + \frac{V(X_{Pv})}{100^2} + \frac{V(X_{Pt})}{50^2}) =$$

$$= \frac{1}{3^2}(\frac{5\lambda}{5^2} + \frac{100\lambda}{100^2} + \frac{50\lambda}{50^2}) = \frac{23}{900}\lambda \approx 0.0256\lambda,$$

$$V(\lambda^*_{Sm}) = V(\frac{X_{Sm} + X_{Pv} + X_{Pt}}{5 + 100 + 50}) = [\text{ober.}] = \frac{V(X_{Sm}) + V(X_{Pv}) + V(X_{Pt})}{(5 + 100 + 50)^2} = \frac{5\lambda + 100\lambda + 50\lambda}{(5 + 100 + 50)^2} =$$

$$= \frac{1}{155}\lambda \approx 0.0065\lambda \text{ respektive}$$

$$V(\lambda^*_{Pt}) = V(\frac{X_{Pv} + X_{Pt}}{100 + 50}) = [\text{ober.}] = \frac{V(X_{Pv}) + V(X_{Pt})}{(100 + 50)^2} = \frac{100\lambda + 50\lambda}{(100 + 50)^2} = \frac{1}{150}\lambda \approx 0.0067\lambda.$$

Alla tre skattningarna är väntevärdesriktiga men Parkvaktens skattning är sämst eftersom den har mycket större varians än de båda andra. Snusmumrikens skattning är bäst. Även om Smusmumrikens mätvärde är osäkert bidrar det till att minska den totala osäkerheten lite grand.

5. (a) Låt X_i vara hallonsyltmängd i för $i = 1, 2, 3$ där $X_i \in N(\mu, \sigma)$ är oberoende.

Eftersom vi vill kunna påvisa om mängden hallonsylt är för liten vill vi testa H_0: $\mu = 1$ mot H_1: $\mu < 1$ med ett ensidigt test på signifikansnivån $\alpha = 0.05$. Det innebär att Filifjonkan inte tänker skämmas förrän någon (Gafsan?) har kunnat övertyga henne, och andra, om att det verkligen finns för lite sylt i hennes rulltårta[1].

Vi har skattningarna $\mu^* = \bar{x} = 0.8667$ respektive

$$\sigma^* = s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} = \sqrt{\frac{0.0139}{3-1}} = 0.0833 \text{ där } \mu^* = \bar{X} \in N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N\left(\mu, \frac{\sigma}{\sqrt{3}}\right)$$

eftersom X_i är oberoende och normalfördelade.

Alt.1: Konfidensintervall. Eftersom vi vill undersöka om μ är mindre än 1 skall vi göra ett intervall som innehåller de tänkbara små värdena för att kunna avgöra om det största av de små värdena är mindre än 1, dvs ett uppåt begränsat intervall. Eftersom $\mu^* = \bar{X} \in N\left(\mu, \sigma/\sqrt{3}\right)$, där σ skattas med s, ges ett ensidigt uppåt begränsat 95 % konfidensintervall för μ av

$$I_\mu = (-\infty,\ \mu^* + t_\alpha(n-1)\cdot\frac{s}{\sqrt{3}}) = (-\infty,\ 0.8667 + \underbrace{t_{0.05}(3-1)}_{2.92}\cdot\frac{0.0833}{\sqrt{3}}) = (-\infty,\ 1.007).$$

Eftersom 1 ligger i intervallet kan H_0: $\mu = 1$ inte förkastas. Mängden hallonsylt är alltså inte signifikant för liten (men bra nära).

Alt.2: Teststorhet. Om H_0: $\mu = 1$ är sann så gäller att $\mu^* \in N\left(1, \sigma/\sqrt{3}\right)$, och därmed att $\frac{\mu^* - 1}{\sigma/\sqrt{3}} \in N(0,\ 1)$ och, när σ skattas med s, att teststorheten $t = \frac{\mu^* - 1}{s/\sqrt{3}} \in t(n-1)$.

Eftersom $t = \frac{\mu^* - 1}{s/\sqrt{3}} = \frac{0.8667 - 1}{0.0833/\sqrt{3}} = -2.7735 \not< -t_\alpha(n-1) = -t_{0.05}(2) = -2.92$ kan H_0 inte förkastas. Mängden hallonsylt är alltså inte signifikant för liten.

(b) Enligt uppgift (a) skall vi förkasta H_0: $\mu = 1$ mot H_1: $\mu < 1$ om $\frac{\bar{x} - 1}{s/\sqrt{n}} < -t_{0.05}(n-1)$, om vi använder teststorheten, eller, om vi använder konfidensintervallet, om $\bar{x} + t_{0.05}(n-1)\cdot\frac{s}{\sqrt{n}} < 1$, vilket är samma sak. Om vi vet att $\sigma = 0.04$ ska vi förkasta H_0 om $\frac{\bar{x} - 1}{0.04/\sqrt{n}} < -\lambda_{0.05}$ istället, där $n = 2$.

Vi vill alltså hitta de värden på μ för vilka vi har styrkan

$$h(\mu) = P(\frac{\bar{X} - 1}{0.04/\sqrt{2}} < -\lambda_{0.05} \parallel \text{om } \bar{X} \in N\left(\mu, \frac{0.04}{\sqrt{2}}\right)) \geq 0.90.$$

[1] Om Filifjonkan är verkligt paranoid kan hon tänkas skämmas tills någon lyckas övertyga henne om att det faktiskt finns tillräckligt med sylt i rulltårtan. I så fall blir hypoteserna istället H_0: $\mu = 1$ och H_1: $\mu > 1$, konfidensintervallet nedåt begränsat och teststorheten ska jämföras med $t_{0.05}(2)$. Eftersom $u = -2.77 \not> 2.92$ kan H_0 inte förkastas. Det finns alltså inte signifikant ”för mycket” sylt i rulltårtan.

Eller så vill hon att det ska vara perfektionistiskt med precis 1 tsk sylt. Då testar hon H_0: $\mu = 1$ mot H_1: $\mu \neq 1$ med ett tvåsidigt intervall och teststorheten ska jämföras med $t_{0.025}(2)$. Eftersom $|t| = 2.77 \not> 4.30$ kan H_0 inte förkastas. Det är alltså inte signifikant ”fel” mängd hallonsylt i rulltårtan.

Eftersom $\mathsf{P}(\dfrac{\bar{X}-1}{0.04/\sqrt{2}} < -\lambda_{0.05} \parallel \text{om } \bar{X} \in N\left(\mu, \dfrac{0.04}{\sqrt{2}}\right)) =$

$= \mathsf{P}(\bar{X} < 1 - \lambda_{0.05} \cdot \dfrac{0.04}{\sqrt{2}} \parallel \text{om } \bar{X} \in N\left(\mu, \dfrac{0.04}{\sqrt{2}}\right)) =$

$= \mathsf{P}(\dfrac{\bar{X}-\mu}{0.04/\sqrt{2}} < \dfrac{1 - \lambda_{0.05} \cdot \frac{0.04}{\sqrt{2}} - \mu}{0.04/\sqrt{2}} \parallel \text{om } \bar{X} \in N\left(\mu, \dfrac{0.04}{\sqrt{2}}\right)) =$

$= \Phi(\dfrac{1 - \lambda_{0.05} \cdot \frac{0.04}{\sqrt{2}} - \mu}{0.04/\sqrt{2}}) = \Phi(\dfrac{1-\mu}{0.04/\sqrt{2}} - \lambda_{0.05}) \geq 0.90$ får vi, med hjälp av $\Phi(\lambda_\alpha) = 1 - \alpha$, att

$\dfrac{1-\mu}{0.04/\sqrt{2}} - \lambda_{0.05} \geq \lambda_{1-0.90} = \lambda_{0.1}$ och därmed att

$\mu \leq 1 - (\lambda_{0.1} + \lambda_{0.05}) \cdot \dfrac{0.04}{\sqrt{2}} = 1 - (1.28 + 1.64) \cdot \dfrac{0.04}{\sqrt{2}} = 0.9172.$

Villkoret är alltså uppfyllt för alla $\mu \leq 0.9172$ tsk.

6. (a) Vi har att $\mathsf{E}(U) = \mathsf{E}(X) + \mathsf{E}(YT) =$ [eftersom Y och T är oberoende] $= \mathsf{E}(X) + \mathsf{E}(Y) \cdot \mathsf{E}(T) =$ $= 2 + 5 \cdot 1 = 7\,\text{m}^2$.

(b) $\mathsf{P}(T > 1) = \displaystyle\int_1^\infty f_T(x)\,dx = \int_1^\infty e^{-x}\,dx = \left[-e^{-x}\right]_1^\infty = e^{-1} \approx 0.37.$

(c) Eftersom T är kontinuerlig börjar vi med att beräkna den betingade fördelningsfunktionen:[2]

$F_{T|T>1}(t) = \mathsf{P}(T \leq t \mid T > 1) = \left[\text{enligt definitionen av betingad sannolikhet}\right] =$

$= \dfrac{\mathsf{P}(T \leq t \cap T > 1)}{\mathsf{P}(T > 1)} = \dfrac{\mathsf{P}(1 < T \leq t)}{\mathsf{P}(T > 1)} = \dfrac{\mathsf{P}(T \leq t) - \mathsf{P}(T \leq 1)}{\mathsf{P}(T > 1)} = \dfrac{F_T(t) - F_T(1)}{\mathsf{P}(T > 1)}$ så att

$f_{T|T>1}(t) = \dfrac{d}{dt} F_{T|T>1}(t) = \dfrac{f_T(t) - 0}{\mathsf{P}(T > 1)} = \dfrac{e^{-t}}{e^{-1}} = e^{-t+1}$ för $t \geq 1$.

Det betingade väntevärdet för hur lång tid mårran sitter ner, givet att hon sitter i mer än en timme, ges av då av $\mathsf{E}(T \mid T > 1) = \displaystyle\int_{-\infty}^{\infty} x \cdot f_{T|T>1}(x)\,dx = \int_1^\infty x \cdot e^{-x+1}\,dx = \left[\text{partialintegrera}\right] =$

$= \left[-x\,e^{-x+1}\right]_1^\infty + \displaystyle\int_1^\infty e^{-x+1}\,dx = e^{-1+1} + \left[-e^{-x+1}\right]_1^\infty = 1 + e^{-1+1} = 2\,\text{h}.$

(d) Eftersom både X och Y är oberoende av T har vi att $\mathsf{E}(X \mid T > 1) = \mathsf{E}(X)$ och $\mathsf{E}(Y \mid T > 1) = \mathsf{E}(Y)$ så att $\mathsf{E}(A \mid T > 1) = \mathsf{E}(X) + \mathsf{E}(Y) \cdot (\mathsf{E}(T \mid T > 1) - 1) = 2 + 5 \cdot (2 - 1) = 7\,\text{m}^2$.

(e) $\mathsf{E}(A) = \mathsf{E}(A \mid T > 1) \cdot \mathsf{P}(T > 1) + \mathsf{E}(A \mid T \leq 1) \cdot \mathsf{P}(T \leq 1) = 7 \cdot e^{-1} + 0 \cdot (1 - e^{-1}) \approx 2.58\,\text{m}^2$.

[2] *Anmärkning*: Man kan också utnyttja att exponentialfördelningen saknar minne, så att den återstående tiden $T_1 = T - 1$, utöver den timme hon redan suttit ner, också är $Exp(1)$-fördelad. Det ger

$F_{T|T>1}(t) = \mathsf{P}(T \leq t \mid T > 1) = \mathsf{P}(T_1 \leq t - 1 \mid T > 1) = \displaystyle\int_{-\infty}^{t-1} f_T(x)\,dx = \int_0^{t-1} e^{-x}\,dx = \left[-e^{-x}\right]_0^{t-1} = 1 - e^{-t+1}$ och

$f_{T|T>1}(t) = \dfrac{d}{dt} F_{T|T>1}(t) = e^{-t+1}$ för $t \geq 1$. Vi får också $\mathsf{E}(T \mid T > 1) = 1 + \mathsf{E}(T_1 \mid T > 1) = 1 + 1 = 2\,\text{h}$.