STAT 251: Homework 6

John Peca-Medlin†
May 18, 2006

1. (a) We are given that U is a standard uniform random variable on $[0,1]$ and it divides this interval into two segments, S, L, with S the length of the shorter segment, L the length of the longer one, and $R = S/L$ the ratio of these length. Then we want to show for $0 \leq r \leq 1$, $F_R(r) := P(R \leq r) = 2r/(1+r)$. Note for U a standard uniform random variable on $[0,1]$ then $f_U(x) = 1/(1-0) = 1$, while for U a random variable over $[0,1/2]$ or $(1/2,1]$, we have $f_U(x) = 1/(1/2) = 2$:

$$\begin{aligned}
F_R(r) &= P(R \leq r) = P(S/L \leq r) = P(S \leq rL) \\
&= P\left(0 \leq U \leq \frac{1}{2}\right) P(U \leq r(1-U)) + P\left(1 \geq U > \frac{1}{2}\right) P(1-U \leq rU) \\
&= P\left(0 \leq U \leq \frac{1}{2}\right) P\left(U \leq \frac{r}{r+1}\right) + P\left(1 \geq U > \frac{1}{2}\right) P\left(\frac{1}{r+1} \leq U\right) \\
&= \int_0^{\frac{1}{2}} f_U(x)\,\mathrm{d}x \int_0^{\frac{r}{r+1}} f_U(x)\,\mathrm{d}x + \int_{\frac{1}{2}}^1 f_U(x)\,\mathrm{d}x \int_{\frac{1}{r+1}}^1 f_U(x)\,\mathrm{d}x \\
&= \frac{1}{2}(2) \int_0^{\frac{r}{r+1}} \mathrm{d}x + \frac{1}{2}(2) \int_{\frac{1}{r+1}}^1 \mathrm{d}x \\
&= (1)\,[x]_0^{\frac{r}{r+1}} + (1)[x]_{\frac{1}{r+1}}^1 \\
&= \frac{r}{r+1} + \left(1 - \frac{1}{r+1}\right) = \frac{2r}{r+1}
\end{aligned}$$

(b) To find the density function f_R of R, note $f_R(r) = \frac{d}{dr} F_R(r)$. Hence:

$$f_R(r) = \frac{d}{dr} F_R(r) = \frac{d}{dr}\left(\frac{2r}{r+1}\right) = 2\frac{d}{dr}\left(1 - \frac{1}{r+1}\right) = \frac{2}{(r+1)^2}$$

(c) Now we want to find $E[R]$ of R:

$$\begin{aligned}
E[R] &= \int_{-\infty}^{\infty} r\, f_R(r)\,\mathrm{d}r = 2\int_0^1 \frac{r}{(r+1)^2}\,\mathrm{d}r \qquad (\text{let } s = r+1) \\
&= 2\int_1^2 \frac{s-1}{s^2}\,\mathrm{d}s = 2\int_1^2 \left(\frac{1}{s} - \frac{1}{s^2}\right)\mathrm{d}s \\
&= 2\left[\ln s + \frac{1}{s}\right]_1^2 = 2\ln 2 - 1 \approx .386294
\end{aligned}$$

(d) And now we want the standard deviation $\mathrm{SD}(R)$ of R, for which we first need $E[R^2]$:

†continues to be known as "*The Yellow Dart*" in literary circles.

$$
\begin{aligned}
E[R^2] &= \int_{-\infty}^{\infty} r^2 f_R(r)\,\mathrm{d}r = 2\int_0^1 \frac{r^2}{(r+1)^2}\,\mathrm{d}r \qquad \text{(add } 0 = (2r+1) - 2r - 1\text{)} \\
&= 2\int_0^1 \frac{(r+1)^2 - 2r - 1}{(r+1)^2}\,\mathrm{d}r = 2\int_0^1 \left(1 - \frac{2r}{(r+1)^2} - \frac{1}{(r+1)^2}\right)\mathrm{d}r \\
&= 2\int_0^1 \mathrm{d}r - 2\int_0^1 \frac{r}{(r+1)^2}\,\mathrm{d}r - 2\int_0^1 \frac{1}{(r+1)^2}\,\mathrm{d}r \\
&= 2 - 2(2\ln 2 - 1) + \left[\frac{2}{r+1}\right]_0^1 = 3 - 4\ln 2 \approx .227411
\end{aligned}
$$

So now we can calculate $\mathrm{SD}(R)$:

$$
\begin{aligned}
\mathrm{SD}(R) &= \sqrt{E[R^2] - E[R]^2} = \sqrt{(3 - 4\ln 2) - (2\ln 2 - 1)^2} \\
&= \sqrt{2 - 4(\ln 2)^2} \approx .279621
\end{aligned}
$$

(e) For $0 \le p \le 1$, we want to find the p^{th} fractional point of R, i.e. the number r_p such that $P(R \le r_p) = p$.

$$
\begin{aligned}
P(R \le r_p) = F_R(r_p) = p = \frac{2r_p}{r_p + 1} &\Leftrightarrow p\,(1 + r_p) = 2r_p \\
\Leftrightarrow p = (2 - p)r_p &\Leftrightarrow r_p = \frac{p}{2 - p}
\end{aligned}
$$

2. We are given that a product sold seasonally yields a net profit of b dollars for each unit sold and a net loss of ℓ dollars for each unit left unsold at the end of the season. Let X be the random number of units of the product that will be ordered by the customers at a particular department store, which must stock this product in advance, and let f be the continuous probability density distribution model for X.

(a) Suppose the store stocks s units of the product. To see why the store's profit G_s is given by the formula $G_s = sb - (b + \ell)(s - X)I_{\{X \le s\}}$, note if the entire supply is exhausted (i.e. $X > s$), then we have $G_s = sb$. Then note that if $X \le s$, then we have the gain is $bX - (s - X)\ell$ and so we need for G_s to equal this an indicator variable $I_{\{X \le s\}}$ (and we need the added sb in front so that it cancels out with the one already mentioned for $X > s$) such that $(sb - (bX - (s - X)\ell))I_{\{X \le s\}} = (sb + s\ell - bx - x\ell)I_{\{X \le s\}} = (b + \ell)(s - X)I_{\{X \le s\}}$. So this shows precisely that the formula for G_s is indeed $G_s = sb - (b+\ell)(s-X)I_{\{X \le s\}}$, as desired.

(b) In order to maximize $E[G_s]$, first we need to calculate $E[G_s]$:

$$
\begin{aligned}
E[G_s] &= sb - (b + \ell)E\left[(s - X)I_{\{X \le s\}}\right] = sb - (b + \ell)\int_0^s (s - x)f(x)\,\mathrm{d}x \\
&= sb - (b + \ell)\left[s\int_0^s f(x)\,\mathrm{d}x - \int_0^s xf(x)\,\mathrm{d}x\right] \\
&= sb - (b + \ell)\left[sF(s) - \int_0^s xf(x)\,\mathrm{d}x\right] \qquad \text{(Using FToC)} \\
\Rightarrow \frac{d}{ds}E[G_s] &= b - (b + \ell)[(F(s) + sf(s)) - sf(s)] \quad \left(\tfrac{d}{ds}(sF(s)) = F(s) + sf(s)\right) \\
&= b - (b + \ell)F(s)
\end{aligned}
$$

Thus, in order to maximize this, we need to set this to zero:

$$0 = b - (b+\ell)F(s^*) \Leftrightarrow F(s^*) = \frac{b}{b+\ell},$$

which is what we wanted to show.

3. (a) Let G_n be the gain from playing roulette with winning \$1 with probabilty $\frac{18}{38}$ and of losing \$1 with probabilty $\frac{20}{38}$, and we want to calculate $p_n := P(A_n) = P(G_n \geq 0)$. First note that $G_n \geq 0$ only if the number of wins is greater than or equal to $n/2$ for n (independent) spins. Since this is a binomial distribution with parameters $n, p = \frac{18}{38}$, then:

$$p_n = P(G_n \geq 0) = \sum_{k \geq n/2}^{n} \binom{n}{k} \left(\frac{18}{38}\right)^k \left(\frac{20}{38}\right)^{n-k},$$

which is all that was wanted.

(b) To see the normal approximation $\widetilde{p}_n$ to $P(A_n)$, we see:

$$\widetilde{p}_n = P\left(Y \geq k = \frac{n}{2}\right) = P\left(Y > \frac{n}{2} - 0.5\right) \approx 1 - \Phi\left(\frac{\frac{n}{2} - 0.5 - np}{\sqrt{np(1-p)}}\right)$$

(c) Plugging 200 and 2000 in for n, we have:

$$\widetilde{p}_{200} \approx 1 - \Phi\left(\frac{\frac{200}{2} - 0.5 - 200p}{\sqrt{200p(1-p)}}\right) = 1 - \Phi(.67) = .2546$$
$$\widetilde{p}_{2000} \approx 1 - \Phi\left(\frac{\frac{2000}{2} - 0.5 - 2000p}{\sqrt{2000p(1-p)}}\right) = 1 - \Phi(2.33) = .0099$$

So as $n \to \infty$, then $\widetilde{p}_n \to 0$, which makes sense since the chance of losing money is greater than winning money, and so for large n the gain will be getting smaller and smaller (I could have just said something like: by the law of large numbers it is so).

4. Let T be a nonnegative random variable with density function $f(t)$, survival function $S(t) = P(T > t)$, and hazard rate $\lambda(t)$. For $\lambda, \alpha > 0$ constants, STFAE:

(i) $\lambda(t) = \lambda\alpha t^{\alpha-1}$ for all $t \geq 0$;

(ii) $S(t) = e^{-\lambda t^\alpha}$ for all $t \geq 0$;

(iii) $f(t) = \lambda\alpha t^{\alpha-1} e^{-\lambda t^\alpha}$ for all $t \geq 0$;

(a) "(ii)$\Rightarrow$(iii)" Assume $S(t) = e^{-\lambda t^\alpha}$ for all $t \geq 0$. Note since $F(t) = P(T \leq t)$, then $F(t) = 1 - S(t)$, and then note the relationship $f(t) = \frac{d}{dt}F(t)$. Then let $u = -\lambda t^\alpha \Rightarrow \frac{du}{dt} = -\lambda\alpha t^{\alpha-1}$:

$$f(t) = \frac{d}{dt}F(t) = \frac{d}{dt}(1 - S(t)) = -\frac{d}{dt}S(t) = -\frac{d}{dt}e^u = -e^u\frac{du}{dt} = \lambda\alpha t^{\alpha-1}e^{-\lambda t^\alpha}$$

(b) "(iii)$\Rightarrow$(ii) and (i)" Assume $f(t) = \lambda\alpha t^{\alpha-1}e^{-\lambda t^\alpha}$ for all $t \geq 0$. To see "(iii)$\Rightarrow$(ii)", note (using the established relationship from part (a)) $\frac{d}{dt}\int S(t)\,\mathrm{d}t =$

$S(t) = -\int f(t)\,\mathrm{d}t$, and using $u = -\lambda\alpha t^{\alpha} \Rightarrow \mathrm{d}u = -\lambda\alpha t^{\alpha-1}\mathrm{d}t^{\ddagger}$:

$$S(t) = -\int f(t)\,\mathrm{d}t = \int (-\lambda\alpha t^{\alpha-1})e^{-\lambda t^{\alpha}}\,\mathrm{d}t = \int e^{u}\,\mathrm{d}u = e^{u} = e^{-\lambda\alpha t^{\alpha}}$$

Now to see "(iii)⇒(i), note by definition, the hazard rate is $\lambda(t) = f(t)/(1 - F(t)) = f(t)/S(t)$:

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{\lambda\alpha t^{\alpha-1}e^{-\lambda t^{\alpha}}}{e^{-\lambda t^{\alpha}}} = \lambda\alpha t^{\alpha-1},$$

as we wanted.

(c) "(i)⇒(ii)" Assume $\lambda(t) = \lambda\alpha t^{\alpha-1}$ for all $t \geq 0$. Then:

$$\begin{aligned} \lambda(t) = \frac{f(t)}{S(t)} = \frac{-\frac{d}{dt}S(t)}{S(t)} &\Rightarrow -\int \lambda(t)\,\mathrm{d}t = \int \frac{S'(t)}{S(t)}\,\mathrm{d}t \\ \Rightarrow \ln(S(t)) &= -\int \lambda(t)\,\mathrm{d}t = -\int \lambda\alpha t^{\alpha-1}\,\mathrm{d}t = -\lambda t^{\alpha} \\ \Rightarrow S(t) &= e^{-\lambda t^{\alpha}} \end{aligned}$$

(d) A random variable T with a standard Rayleigh distribution has a Weibull distribution. By page 7 of the handout for Lecture 14, we see that T has a hazard rate $\lambda(t) = t$ for $t \geq 0$. Note by part (c) above, we have the relationship $S(t) = \exp\left(-\int \lambda(t)\,\mathrm{d}t\right)$, so:

$$S(t) = \exp\left(-\int \lambda(t)\,\mathrm{d}t\right) = \exp\left(-\int t\,\mathrm{d}t\right) = e^{-t^2/2}$$

Since $S(t) = e^{-\lambda t^{\alpha}}$ for a Weibull(α, λ) distribution, then we see that T has shape parameter $\alpha = 2$ and intensity parameter $\lambda = \frac{1}{2}$. This could have been derived directly from $\lambda(t)$, but I like integral signs so I decided to throw in a few.

(e) Suppose T has a Weibull(α, λ) distribution, and we want to show that $Y = T^{\alpha}$ has an exponential distribution with rate parameter λ. Let $S_T(t)$ be the survival function of T, so $F_Y(y) = 1 - P(Y > y) = 1 - P(T^{\alpha} > y) = 1 - P(T > y^{1/\alpha}) = 1 - S_T(y^{1/\alpha}) = 1 - e^{-\lambda y}$. Since an exponential random variable with parameter λ has a distribution function exactly of this form, we see Weibull$(\alpha, \lambda) \sim$ Exp(λ), as desired. Using the method from class, note $t = g^{-1}(y) = y^{1/\alpha}$ so for $y \geq 0$:

$$\begin{aligned} f_Y(y) = f_T(t)\frac{dt}{dy} &= f_T(y^{1/\alpha})\frac{d}{dy}(y^{1/\alpha}) \\ &= \left(\lambda\alpha(y^{1/\alpha})^{\alpha-1}e^{-\lambda(y^{1/\alpha})^{\alpha}}\right)\left(\frac{y^{1/\alpha-1}}{\alpha}\right) \\ &= \lambda e^{-\lambda y} \end{aligned}$$

This is precisely the density function for an exponential random variable with parameter λ for $y \geq 0$, as we wanted to show (again).

‡note $\int f(t)\,\mathrm{d}t = \int_0^t f(s)\,\mathrm{d}s$, I just used the former for convenience sake since for the integrals used here, the functions evaluated at 0 are 0.

(f) Let $Y = T^k \sim \text{Exp}(\lambda) \Rightarrow X = \lambda Y \sim \text{Exp}(1)$, so $T = Y^{1/\alpha} = \left(\frac{X}{\lambda}\right)^{1/\alpha}$:

$$E[T^k] = \frac{E[X^{k/\alpha}]}{\lambda^{k/\alpha}} = \frac{1}{\lambda^{k/\alpha}} \int_0^\infty x^{k/\alpha} e^{-x}\,\mathrm{d}x = \frac{\Gamma\left(1+\frac{k}{\alpha}\right)}{\lambda^{k/\alpha}} = \frac{k\Gamma\left(\frac{k}{\alpha}\right)}{\alpha\lambda^{k/\alpha}}$$

Thus, for $k = 1, 2$ we can get $E[T]$ and $E[T^2]$:

$$E[T] = \frac{\Gamma\left(1+\frac{1}{\alpha}\right)}{\lambda^{1/\alpha}} = \frac{\Gamma\left(\frac{1}{\alpha}\right)}{\alpha\lambda^{1/\alpha}}$$
$$E[T^2] = \frac{\Gamma\left(1+\frac{2}{\alpha}\right)}{\lambda^{2/\alpha}} = \frac{2\Gamma\left(\frac{2}{\alpha}\right)}{\alpha\lambda^{2/\alpha}}$$

$$\Rightarrow \text{SD}(T) = \sqrt{E[T^2] - E[T]^2} = \frac{1}{\lambda^{1/\alpha}}\sqrt{\Gamma\left(1+\frac{2}{\alpha}\right) - \left(\Gamma\left(1+\frac{1}{\alpha}\right)\right)^2}$$
$$= \frac{1}{\alpha\lambda^{1/\alpha}}\sqrt{2\alpha\Gamma\left(\frac{2}{\alpha}\right) - \left(\Gamma\left(\frac{1}{\alpha}\right)\right)^2}$$

The median m of T would have the property $P(T < m) = P(T \geq m) = 1/2$. Since for a continuous random variable $P(T \geq m) = P(T > m) = S(m)$, and for a Weibull distribution, $S(t) = \exp(-\lambda t^\alpha)$, so:

$$S(m) = \frac{1}{2} = e^{-\lambda m^\alpha} \Rightarrow -\lambda m^\alpha = \ln(1/2) = -\ln 2$$
$$\Rightarrow m = \left(\frac{\ln 2}{\lambda}\right)^{1/\alpha}$$

(g) Suppose $T \sim \text{Weibull}(\alpha, \lambda)$ for $\alpha = 1/3$ and $\lambda = 2$. So we just need to plug these values into the formulas of part (f):

$$E[T] = \frac{\Gamma(4)}{(2)^3} = \frac{4!}{8} = 3$$
$$E[T^2] = \frac{\Gamma(7)}{(2)^6} = \frac{7!}{64} = \frac{315}{4}$$
$$\Rightarrow \text{SD}(T) = \sqrt{E[T^2] - E[T]^2} = \sqrt{\left(\frac{315}{4}\right) - (3)^2} = \frac{3\sqrt{31}}{2} = 8.35165$$
$$m = \left(\frac{\ln 2}{2}\right)^3 = .041628$$
$$P(T \geq E[T]) = S(E[T]) = S(3) = e^{-2\cdot 3^{1/3}} = .055883$$

5. (a) An urn contains ν balls, labeled from 1 to ν. Balls will be drawn at random with replacement until the first time W_ν that each ball has been seen at least once. For $n = 1, \ldots, \nu$ let $N_{\nu,n}$ be the number of balls that aren't drawn at least once during the first n draws. We want to express the event $\{W_\nu \leq n\}$ in terms of $N_{\nu,n}$. Note for $\{W_\nu \leq n\}$ to occur, then this is the same as the event $\{N_{\nu,n} = 0\}$, since we need that there be no balls left that have occurred by the n^{th}which is what we wanted.

(b) Let $F_{i;\nu,n}$ be the event ball i does not appear in the first n draws. It is then easy to see $N_{\nu,n} = \sum_{i=1}^{n} I_{\{F_{i;\nu,n}\}}$ since the right side (using indicator variables) would be precisely the number of balls that have not appeared by the first n draws, which is what $N_{\nu,n}$ is. Next note:

$$\begin{aligned}
P\left[\frac{W_\nu - \nu\log(\nu)}{\nu} \leq x\right] &= P[W_\nu \leq x\nu + \nu\log(\nu)] = P[W_\nu \leq n] = P[N_{\nu,n} = 0] \\
&= P\left[\bigcap_{i=1}^{\nu} F_{i;\nu,n}^c\right] = 1 - P\left[\bigcup_{i=1}^{\nu} F_{i;\nu,n}\right] \\
&= 1 - \sum_{k=1}^{\nu} \binom{j}{k} (-1)^{k-1} \left(1 - \frac{k}{j}\right)^n \\
&= 1 - \sum_{k=1}^{\infty} \binom{\nu}{k} (-1)^{k-1} \left(1 - \frac{k}{\nu}\right)^{\nu(x+\log(\nu))}
\end{aligned}$$

Next, looking at the part inside the infinite sum, notice the following relationships: for ν large, $(\nu)_k \sim \nu^k$, and also