Math 5010 § 1. **Second Midterm Exam** Name: SomeSolutions
Treibergs $\sigma\tau$ April 27, 2009

Solutions to the Extra Sample Problems from the Last Third of the Class.

1. *Let X_n be a sequence of independent exponential variables with parameter λ for all n. Show that the sample means converge in probability as $n \to \infty$,*

$$\frac{1}{n}(X_1 + \cdots + X_n) \xrightarrow{P} \frac{1}{\lambda}.$$

Let $S_n = X_1 + \cdots + X_n$ as usual and observe that because of independence and the behavior of variance under multiplication by a constant, the random variable S_n/n satisfies

$$\mathbf{E}\left(\frac{1}{n}S_n\right) = \frac{1}{n}\,\mathbf{E}\left(\sum_{j=1}^{n} X_j\right) = \frac{1}{n}\sum_{j=1}^{n}\mathbf{E}(X_j) = \frac{1}{n}\sum_{j=1}^{n}\frac{1}{\lambda} = \frac{1}{\lambda};$$

$$\mathbf{Var}\left(\frac{1}{n}S_n\right) = \frac{1}{n^2}\,\mathbf{Var}\left(\sum_{j=1}^{n} X_j\right) = \frac{1}{n^2}\sum_{j=1}^{n}\mathbf{Var}(X_j) = \frac{1}{n^2}\sum_{j=1}^{n}\frac{1}{\lambda^2} = \frac{1}{n\lambda^2}.$$

Recall the version of Chebychov's inequality, gotten by applying Theorem 7.4.16.iii to $h(x) = (x - \mathbf{E}(X))^2$. For all $a > 0$,

$$\mathbf{P}(|X - \mathbf{E}(X)| \ge a) \le \frac{\mathbf{Var}(X)}{a^2}. \tag{1}$$

Recall also the meaning of $Z_n \xrightarrow{P} Z$ "convergence in probability." It means for every $\varepsilon > 0$,

$$\lim_{n\to\infty}\mathbf{P}\left(|Z_n - Z| \ge \varepsilon\right) = 0.$$

In our case, $Z_n = S_n/n$, $Z = 1/\lambda$ and apply (1) with $a = \epsilon$, where $\epsilon > 0$ is any positive number. Then

$$\mathbf{P}\left(\left|\frac{1}{n}S_n - \frac{1}{\lambda}\right| \ge \varepsilon\right) \le \frac{\mathbf{Var}\left(\frac{1}{n}S_n\right)}{\varepsilon^2} = \frac{1}{\varepsilon^2\,\lambda^2\,n} \to 0 \text{ as } n \to \infty.$$

Thus the sample average converges in probability to its mean, as to be shown.

2. *Let P and Q be two points chosen independently and uniformly on $(0,1)$. Let D denote the distance between the two points. Find the probability density function $f_D(d)$.*

We observe that the distance must satisfy $0 < D < 1$. Suppose that the coordinates of the two points are $x, y \in I = (0,1)$. From independence, the joint density is $f(x,y) = f_X(x)f_Y(y) = 1$ if $0 < x, y < 1$ and $f(x,y) = 0$ otherwise. The event $A_d = \{D \le d\} = \{(x,y) \in I^2 : |x-y| \le d\} = \{(x,y) \in I^2 : x - d \le y \le x + d\}$ is the band of vertical width $2d$ centered on the $x = y$ line in the unit square. Its probability is for $0 < d < 1$,

$$F_D(d) = \mathbf{P}(D \le d) = \int_{A_d} f(x,y)\,dx\,dy = 1 - 2\int_d^1\int_0^{x-d} dy\,dx = 2d - d^2$$

where we have subtracted the upper and lower triangle from the unit square to find the integral over the band. Thus

$$f_D(d) = F_D'(d) = \begin{cases} 0, & \text{if } d \le 0; \\ 2 - 2d, & \text{if } 0 < d < 1; \\ 0, & \text{if } 1 \le d. \end{cases}$$

3. *A system consisting of one original unit plus a spare can function for a random amount of time X. If the density of X is given (in months) by*

$$f_X(x) = \begin{cases} Cxe^{-x/2}, & \text{if } x > 0; \\ 0, & \text{if } x \le 0. \end{cases}$$

   (a) *Find C to make $f_X(x)$ a pdf.*
   (b) *What is the probability that the system functions for at least five months?*
   (c) *Find the moment generating function $M_X(t)$. For what t is it defined?*
   (d) *Find $\mathbf{E}(X)$ and $\mathbf{Var}(X)$.*

For f_X to be a pdf, its total probability has to be one. Integrating by parts, let $u = x$ and $dv = e^{-x/2}$ so $du = dx$ and $v = -2e^{-x/2}$ yielding

$$1 = C\int_0^\infty xe^{-x/2}\,dx = \left(-2Cx\,e^{-x/2}\right)\Big|_0^\infty + 2C\int_0^\infty e^{-x/2}\,dx = 0 + \left(-4C\,e^{-x/2}\right)\Big|_0^\infty = 4C$$

so $C = 1/4$. Hence the probability that the system lasts at least five months is

$$\mathbf{P}(X \ge 5) = \frac{1}{4}\int_5^\infty x\,e^{-x/2}dx = \left(-\frac{1}{2}x\,e^{-x/2} - e^{-x/2}\right)\Big|_5^\infty = \frac{7}{2}\,e^{-5/2} \approx 0.287.$$

The moment generating function is $M_X(t) = \mathbf{E}(e^{tX}) =$

$$\frac{1}{4}\int_0^\infty e^{tx}\,x\,e^{-\frac{x}{2}}\,dx = \frac{1}{4}\int_0^\infty x\,e^{(t-\frac{1}{2})x}dx = \left(\frac{xe^{(t-\frac{1}{2})x}}{4(t-\frac{1}{2})} - \frac{e^{(t-\frac{1}{2})x}}{4(t-\frac{1}{2})^2}\right)\Bigg|_0^\infty = \frac{1}{(1-2t)^2}$$

where $t < 1/2$ to make sure the integral converges at infinity. Because $M_X(0) = 1$, the expectation and variance of X are

$$\mathbf{E}(X) = \frac{M_X'(0)}{M_X(0)} = \frac{d}{dt}\Big|_{t=0}\log M_X(t) = \frac{4}{1-2t}\Big|_{t=0} = 4;$$
$$\mathbf{Var}(X) = \frac{M_X''(0)}{M_X(0)} - \frac{M_X'(0)^2}{M_X(0)^2} = \frac{d^2}{dt^2}\Big|_{t=0}\log M_X(t) = \frac{8}{(1-2t)^2}\Big|_{t=0} = 8.$$

As a double check that these values are correct, this variable is distributed according to the gamma density with parameters $\lambda = 1/2$ and $r = 2$, as on p. 320.

4. *Suppose $X_1, X_2, X_3, \ldots$ is a sequence of independent random variables all exponentially distributed with parameter λ. Let*

$$Y_n = \frac{S_n - \mathbf{E}(S_n)}{\sqrt{\mathbf{Var}(S_n)}}$$

where $S_n = X_1 + X_2 + \cdots + X_n$. Show that

$$Y_n \xrightarrow{D} Z \qquad \text{as} \qquad n \to \infty$$

converges in distribution, where $Z \sim N(0,1)$ is the standard normal variable. (Don't quote CLT.)

The method is the same as proving the de Moivre-Laplace Theorem, or the Central Limit Theorem. By the Continuity Theorem, it suffices to show that for some $b > 0$, the moment generating functions are finite for $|t| < b$ and converge: for all $|t| \le b/2$,

$$M_{X_n}(t) \to M_Z(t) \qquad \text{as} \qquad n \to \infty.$$

We know that $\mathbf{E}(X_n) = 1/\lambda$ so that $\mathbf{E}(S_n) = n/\lambda$. Also $\mathbf{Var}(X_n) = 1/\lambda^2$ so that by independence, $\mathbf{Var}(S_n) = n/\lambda^2$. Hence

$$Y_n = \frac{\lambda S_n}{\sqrt{n}} - \sqrt{n}.$$

The moment generating function for the exponential rv, *e.g.,* $X_1 \sim \text{exponential}(\lambda)$ is $M_{X_1}(t) = \lambda/(\lambda - t)$ as long as $t < \lambda$. Because the mgf of a sum of independent variables is a product of mgf's, the moment generating function for Y_n is

$$\begin{aligned}
M_{Y_n}(t) &= \mathbf{E}\left(e^{tY_n}\right) \\
&= \mathbf{E}\left(\exp\left(\frac{t\lambda S_n}{\sqrt{n}} - t\sqrt{n}\right)\right) \\
&= \exp\left(-t\sqrt{n}\right)\mathbf{E}\left(\exp\left(\frac{t\lambda}{\sqrt{n}}\, S_n\right)\right) \\
&= \exp\left(-t\sqrt{n}\right)\left(\frac{\lambda}{\lambda - \frac{\lambda t}{\sqrt{n}}}\right)^n
\end{aligned}$$

since we have viewed the "t" in the mgf as $t\lambda/\sqrt{n}$ which is less than λ for large n. Taking logarithms, and expanding the logarithm we see that for, say $|t| \le 1$, as $n \to \infty$,

$$\begin{aligned}
\log M_{Y_n}(t) &= -t\sqrt{n} - n\,\log\left(1 - \frac{t}{\sqrt{n}}\right) \\
&= -t\sqrt{n} + n\left[\frac{t}{\sqrt{n}} + \frac{t^2}{2n} + \mathbf{O}\left(\frac{1}{n^{\frac{3}{2}}}\right)\right] \\
&= \frac{t^2}{2} + \mathbf{O}\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

It follows from the continuity of the exponential that for all $|t| \le 1/2$,

$$\lim_{n\to\infty} M_{Y_n}(t) = e^{\frac{1}{2}t^2} = M_Z(t)$$

which completes the argument.

5. *Suppose X and Y are independent standard normal variables. Find the pdf for $Z = X^2+Y^2$.*

By independence, the joint density is

$$f(x,y) = f_X(x)f_Y(y) = \left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\right)\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y^2}\right) = \frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}.$$

By changing to polar coordinates $\rho = \sqrt{x^2+y^2}$ and $\theta = \operatorname{Atn}(y/x)$, we find that the cumulative density is for $0 \le z$,

$$\begin{aligned}
F_Z(z) &= \mathbf{P}(Z \le z) = \mathbf{P}(X^2+Y^2 \le z) = \mathbf{P}\left(\sqrt{X^2+Y^2} \le \sqrt{z}\right) \\
&= \frac{1}{2\pi}\iint_{\sqrt{x^2+y^2}\le\sqrt{z}} e^{-\frac{1}{2}(x^2+y^2)}\,dx\,dy = \frac{1}{2\pi}\int_{\theta=0}^{2\pi}\int_{\rho=0}^{\sqrt{z}} e^{-\frac{1}{2}\rho^2}\,\rho\,d\rho\,d\theta \\
&= \left(-e^{-\frac{1}{2}\rho^2}\right)\Big|_0^{\sqrt{z}} = 1 - e^{-\frac{1}{2}z}
\end{aligned}$$

Hence $f_Z(z) = F_Z'(z) = \frac{1}{2}e^{-z/2}$ for $z > 0$, which is exponential with parameter $\lambda = \frac{1}{2}$.

6. *Suppose X and Y are independent standard normal variables. Find* $\mathbf{E}(\max(|X|,|Y|))$.

   The joint density $f(x,y)$ is as in Problem 5. Note that, the function $g(x,y) = \max(|x|,|y|)$ is even with respect to reflections across the axes and across the $x = y$ line

$$g(x,y) = g(x,-y) = g(-x,y) = g(y,x)$$

   so that knowing $g(x,y) = x$ for the sector $0 \le y \le x$ gives the function on the whole plane by reflection. $f(x,y)$ has the same symmetries, thus the expectation can be computed on each of the eight sectors. Switching to polar coordinates, using $x = r\cos\theta$,

$$\begin{aligned}
\mathbf{E}(g(X,Y)) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f(x,y)\, dx\, dy \\
&= \frac{8}{2\pi}\int_0^{\infty}\int_0^{x} x e^{-\frac{1}{2}(x^2+y^2)} dy\, dx \\
&= \frac{4}{\pi}\int_0^{\infty}\int_0^{\frac{\pi}{4}} r\cos\theta\, e^{-\frac{1}{2}r^2} r\ d\theta\, dr \\
&= \frac{4}{\pi}\int_0^{\infty} \Big(\sin\theta\Big)\Big|_0^{\frac{\pi}{4}}\ e^{-\frac{1}{2}r^2}\ r^2\, dr \\
&= \frac{2\sqrt{2}}{\pi}\int_0^{\infty} e^{-\frac{1}{2}r^2}\ r^2\, dr.
\end{aligned}$$

   You can recognize this as an integral for variance of a normal variable. Otherwise, integrating by parts with $u = r$ so $du = dr$ and $dv = r\, e^{-r^2/2}\, dr$ so $v = -e^{r^2/2}\, dr$,

$$\begin{aligned}
\mathbf{E}(g(X,Y)) &= \frac{2\sqrt{2}}{\pi}\left(-r^2\, e^{-\frac{1}{2}r^2}\right)\Big|_0^{\infty} + \frac{2\sqrt{2}}{\pi}\int_0^{\infty} e^{-\frac{1}{2}r^2}\ dr \\
&= \frac{\sqrt{2}}{\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}r^2}\ dr = \frac{2}{\sqrt{\pi}},
\end{aligned}$$

   where we used $\int_{-\infty}^{\infty} e^{-r^2/2}\, dr = \sqrt{2\pi}$ and the evenness of the normal integrand.

7. *Let C be a circle of radius R. Find the average length of a random chord. This problem has many possible answers, and that's why it's called* **Bertrand's Paradox**. *Of course what "random chord" means has not been specified and different notions yield different results.*

   (a) Suppose that we pick two endpoints of a chord at random on the circumference of C independently and uniformly. Then what is the expected chord length?

   (b) Let's suppose we pick a random line according to the kinematic density, the one that is invariant under Euclidean motions of the plane. That means, if h is the perpendicular distance of the line from the origin and θ is its direction, then (h,θ) is uniform in $[0,R]\times[0,2\pi)$. What is the expected chord length now?

   (a.) Instead of choosing two independent angles uniformly $\theta_1, \theta_2 \in [0,2\pi)$, it is simpler to choose $a \in [0,2\pi)$ and $b \in [0,2\pi)$ instead and transform by $(\theta_1,\theta_2) = (a, a+b)$ which has unit Jacobian, so the joint density is constant $f(a,b) = 1/(4\pi^2)$ on $0 \le a, b < 2\pi$. Then use 2π periodicity in angles. The chord length depends on b only, and it is given by

$$x = 2R\sin\left(\frac{b}{2}\right),$$

   because if the circle is rotated so that midpoint of the chord is on the positive x-axis, then the height of the endpoint is half of the chord length. The expectation is thus

$$\mathbf{E}(X) = \frac{2R}{4\pi^2}\int_{a=0}^{2\pi}\int_{b=0}^{2\pi} \sin\left(\frac{b}{2}\right)\, db\, da = \frac{4R}{\pi}.$$

(b.) This time the chord length only depends on h and is given by

$$x = 2\sqrt{R^2 - h^2}.$$

The density is uniform, hence constant $f(h, \theta) = 1/(2\pi R)$ on $0 \le \theta < 2\pi$ and $0 \le h \le R$. The expectation is thus

$$\mathbf{E}(X) = \frac{2}{2\pi R} \int_{h=0}^{R} \int_{\theta=0}^{2\pi} \sqrt{R^2 - h^2}\, d\theta\, dh = \frac{\pi R}{2}.$$

8. *Suppose that a lighthouse is on an island at a point L which is a distance of A miles from a straight shoreline, that O is the closest point on the shoreline to the lighthouse and that its beacon is rotating at a constant velocity. Let Q be the point on shore where the light is pointing and X the signed distance from O. Given that the beacon is pointing toward the shoreline at a random instant, what is the probability density function of the distance X?*

   The formula is $x = A \tan\theta$ where θ is the angle $\angle OLQ$ and we assume that the light can be pointing in direction θ uniformly in $[-\pi, \pi)$. The light is pointing shoreward if $|\theta| < \pi/2$. Thus, using uniformity of θ,

$$F_X(x) = \mathbf{P}(X \le x \mid |\theta| < \frac{\pi}{2}) = \frac{\mathbf{P}(X \le x \text{ and } |\theta| < \frac{\pi}{2})}{\mathbf{P}(|\theta| < \frac{\pi}{2})} = \frac{\mathbf{P}\left(\theta \le \operatorname{Atn}\left(\frac{x}{A}\right) \text{ and } |\theta| < \frac{\pi}{2}\right)}{\frac{1}{2}}$$
$$= 2\mathbf{P}\left(-\frac{\pi}{2} < \theta \le \operatorname{Atn}\left(\frac{x}{A}\right)\right) = 2\frac{\operatorname{Atn}\left(\frac{x}{A}\right) + \frac{\pi}{2}}{2\pi} = \frac{1}{\pi} \operatorname{Atn}\left(\frac{x}{A}\right) + \frac{1}{2}.$$

   It follows that the pdf for X is

$$f_X(x) = F_X'(x) = \frac{1}{A\pi\left(1 + \frac{x^2}{A^2}\right)},$$

   which is a stretched Cauchy distribution.

9. *Let X_n have negative binomial distribution with parameters n and $p = 1 - q$. Show using generating functions that if $n \to \infty$ in such a way that $\lambda = nq$ remains constant, then*

$$\lim_{n\to\infty} \mathbf{P}(X_n = k) = \frac{e^{-\lambda}\lambda^k}{k!}.$$

   The pgf for negative binomial variable X_n is

$$G_n(s) = \left(\frac{ps}{1 - qs}\right)^n.$$

   Substituting $q = \lambda/n$ so that $n = (n - \lambda)/n$, we see that

$$G_n(s) = \left(\frac{\frac{(n-\lambda)s}{n}}{1 - \frac{\lambda s}{n}}\right)^n = \left(\frac{(n-\lambda)s}{n - \lambda s}\right)^n$$

   Taking limits of the logarithm and using l'Hôpital's rule,

$$\lim_{n\to\infty} \log G_n(s) = \lim_{n\to\infty} \frac{\log(n-\lambda) + \log s - \log(n - \lambda s)}{\frac{1}{n}}$$
$$= \lim_{n\to\infty} \frac{\frac{1}{n-\lambda} - \frac{1}{n-\lambda s}}{-\frac{1}{n^2}} = \lim_{n\to\infty} \frac{\frac{(1-s)\lambda}{(n-\lambda)(n-\lambda s)}}{-\frac{1}{n^2}} = (s-1)\lambda.$$

Hence, using the continuity of the exponential,

$$\lim_{n\to\infty}\left(\sum_{k=0}^{\infty}\mathbf{P}(X_n=k)s^k\right)=\lim_{n\to\infty}G_n(s)=e^{(s-1)\lambda}=\sum_{k=0}^{\infty}\frac{e^{-\lambda}\,\lambda^k\,s^k}{k!}$$

Assuming that the limit and sum can be interchanged (which is OK because the sum converges uniformly for $-1\le s\le 1$), and equating coefficients,

$$\lim_{n\to\infty}\mathbf{P}(X_n=k)=\frac{e^{-\lambda}\,\lambda^k}{k!}.$$

10. *Let $X_1, X_2, \ldots, X_n$ be independent random variables uniformly distributed on $[0,1]$. Let $Y=\max\{X_1,X_2,\ldots,X_n\}$. Find the cumulative distribution function, density, expectation and variance of Y. What is the probability that $Y>1-\frac{1}{n}$? What is the limiting probability as $n\to\infty$?*

We use the same idea that you used to solve the homework problem for the maximum of discrete uniform variables. The cumulative density for any X_i is

$$F(x)=\mathbf{P}(X_i\le x)=\begin{cases}0, & \text{if } x\le 0;\\ x, & \text{if } 0<x<1;\\ 1, & \text{if } 1\le x.\end{cases}$$

Then, by independence,

$$\begin{aligned}F_Y(y)=\mathbf{P}(Y\le y)&=\mathbf{P}(X_1\le y \text{ and } X_2\le y \text{ and } \ldots \text{ and } X_n\le y)\\ &=\prod_{i=1}^{n}\mathbf{P}(X_i\le y)=F(y)^n=\begin{cases}0, & \text{if } y\le 0;\\ y^n, & \text{if } 0<y<1;\\ 1, & \text{if } 1\le y.\end{cases}\end{aligned}$$

It follows that

$$f_Y(y)=F_Y'(y)=nF(y)^{n-1}\,f(y)=\begin{cases}n\,y^{n-1}, & \text{if } 0<y<1;\\ 0, & \text{otherwise.}\end{cases}$$

The expectation and second moments are

$$\mathbf{E}(Y)=n\int_0^1 y^n\,dy=\frac{n}{n+1},\qquad \mathbf{E}(Y^2)=n\int_0^1 y^{n+1}\,dy=\frac{n}{n+2}$$

Thus

$$\mathbf{Var}(Y)=\mathbf{E}(Y^2)-\mathbf{E}(Y)^2=\frac{n}{n+2}-\frac{n^2}{(n+1)^2}=\frac{n}{(n+1)^2\,(n+1)}.$$

Finally

$$\mathbf{P}\left(Y>1-\frac{1}{n}\right)=1-F_Y\left(1-\frac{1}{n}\right)=1-\left(1-\frac{1}{n}\right)^n\to 1-e^{-1}$$

as $n\to\infty$.

11. *Two types of coins are produced in a factory, a fair coin, and a biased one that comes up heads 55% of the time. You have one of the coins, but do not know if it is the fair one or the biased one. In order to determine which coin you have, you perform a statistical test: toss the coin 1000 times. If the coin lands on heads 525 or more times, then you conclude that it is the biased one. If it lands heads less than 525 times, then conclude that it is the fair*

coin. If the coin is actually fair, what is the probability that we reach a false conclusion? What would it be if the coin were biased?

The actual probabilities can be computed using the binomial distribution. Since n is large, we can approximate the answer using the normal distribution to get the numerical answer.

The number of heads if the coin is fair is given by X which is distributed according to the binomial distribution with parameters $n = 1000$ and $p = 0.50$. Let Y be the number of heads for the unfair coin, which is binomial with parameters $n = 1000$ and $p = 0.55$. For the fair coin, the false conclusion is reached if it lands heads at least 525 times. For the unfair one, the false conclusion is if there are less that 525 heads. Thus we seek $\mathbf{P}(X \ge 525) = 1 - \mathbf{P}(X < 525)$ and $\mathbf{P}(Y < 525)$.

The de Moivre-Laplace theorem says that standardized binomial variables approach the normal one. The approximation is valid, according to the rule of thumb, if $npq \ge 10$. In our case $npq = 1000(0.5)(0.5) = 250$ or $npq = 1000(.55)(.45) = 247.5$ which is well within the rule of thumb. The probability that the test is wrong for the unfair coin is, by standardizing, using $\mathbf{E}(Y) = np$ and $\mathbf{Var}(Y) = npq$,

$$\mathbf{P}(Y < 525) = \mathbf{P}\left(\frac{Y - np}{\sqrt{npq}} < \frac{525 - np}{\sqrt{npq}}\right) \approx \mathbf{P}\left(Z < \frac{525 - 550}{\sqrt{247.5}}\right) = \Phi(-1.589) \approx 0.0560.$$

where $Z \sim N(0,1)$ is standard normal and $\Phi(z)$ its cdf. Similarly

$$\mathbf{P}(X < 525) = \mathbf{P}\left(\frac{Y - np}{\sqrt{npq}} < \frac{525 - np}{\sqrt{npq}}\right) \approx \mathbf{P}\left(Z < \frac{525 - 500}{\sqrt{250.0}}\right) = \Phi(1.581) \approx 0.9430.$$

Hence, for the fair coin, the test gives the wrong answer with $\mathbf{P}(X \ge 525) \approx 1 - 0.9430 = 0.0570$. The normal values were found by interpolating in a table of cumulative normals.

12. *Let U be a uniformly distributed on $(0,1)$ and $r > 0$. Show how to use U to simulate the distribution*

$$F_X(x) = \begin{cases} 0, & \text{if } x \le 1; \\ 1 - x^{-r}, & \text{if } 1 < x. \end{cases}$$

We are looking for a function $g : (0,1) \to (1,\infty)$ so that $X = g(U)$ has the desired distribution. Then for $x > 0$, since F_X is strictly increasing

$$1 - x^{-r} = F_X(x) = \mathbf{P}(X \le x) = \mathbf{P}(g(U) \le x) = \mathbf{P}(U \le g^{-1}(x)) = F_U(g^{-1}(x)) = g^{-1}(x)$$

since U is uniform ($F_U(u) = u$ for $0 < u < 1$.) Solving for $x = g(u)$,

$$1 - g(u)^{-r} = u \qquad \Longrightarrow \qquad g(u) = (1-u)^{-\frac{1}{r}}$$

13. *Let U and V be uniformly distributed on $(0,1)$. Find the cumulative distribution function and probability density function of $Z = UV$.*

For $0 < z < 1$, the region $A_z = \{(u,v) \in I^2 : uv \le z\}$ is bounded by the coordinate axes, the lines $x = 1$ and $y = 1$ and the curve $y = z/x$ from $x = z$ to $x = 1$. The density $f(u,v) = 1$ for $0 < u, v < 1$. Hence for $0 < z \le 1$,

$$F_Z(z) = \mathbf{P}((U,V) \in A_z) = \int_{\xi=0}^{z} \int_{\eta=0}^{1} d\eta\, d\xi + \int_{\xi=z}^{1} \int_{\eta=0}^{z/\xi} d\eta\, d\xi = z + \int_{\xi=z}^{1} \frac{z}{\xi}\, d\xi = z - z\log z.$$

Hence for $0 < z \le 1$,

$$f_Z(z) = F_Z'(z) = -\log z.$$

14. *Let $0 < \sigma, \tau$ and $|\rho| < 1$ be constants. Suppose X and Y be jointly distributed satisfying the bivariate normal density for $(x, y) \in \mathbf{R}^2$,*

$$f(x,y) = \frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} e^{-\frac{1}{2}Q(x,y)}$$

where the quadratic form is given by

$$Q(x,y) = \frac{1}{1-\rho^2}\left\{\frac{x^2}{\sigma^2} - \frac{2\rho xy}{\sigma\tau} + \frac{y^2}{\tau^2}\right\}.$$

Find the marginal density $f_Y(y)$ and the correlation coefficient of X and Y.

By completing the square, we obtain

$$Q(x,y) = \frac{1}{1-\rho^2}\left\{\frac{x^2}{\sigma^2} - \frac{2\rho xy}{\sigma\tau} + \frac{\rho^2 y^2}{\tau^2} + \frac{(1-\rho^2)y^2}{\tau^2}\right\} = \frac{1}{(1-\rho^2)\sigma^2}\left(x - \frac{\rho\sigma y}{\tau}\right)^2 + \frac{y^2}{\tau^2}.$$

By the formula for marginal density

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)\,dx \\
&= \frac{1}{2\pi\sigma\tau\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)\sigma^2}\left(x - \frac{\rho\sigma y}{\tau}\right)^2 - \frac{y^2}{2\tau^2}}\,dx \\
&= \frac{1}{\tau\sqrt{2\pi}} e^{-\frac{y^2}{2\tau^2}} \left(\frac{1}{\sigma\sqrt{2\pi(1-\rho^2)}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)\sigma^2}\left(x - \frac{\rho\sigma y}{\tau}\right)^2}\,dx\right) \\
&= \frac{1}{\tau\sqrt{2\pi}} e^{-\frac{y^2}{2\tau^2}}
\end{aligned}$$

because, for fixed y, the integral is the total probability of the density of the variable $Z = X - \frac{\rho\sigma y}{\tau}$ is a $Z \sim N\left(\frac{\rho\sigma y}{\tau}, (1-\rho^2)\sigma^2\right)$ normal variable, whose density is

$$f_Z(x) = \frac{1}{\sigma\sqrt{2\pi(1-\rho^2)}} e^{-\frac{1}{2(1-\rho^2)\sigma^2}\left(x - \frac{\rho\sigma y}{\tau}\right)^2}$$

It follows that $Y \sim N(0, \tau^2)$. By symmetry,

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

so that $X \sim N(0, \sigma^2)$. Because $\mathbf{E}(X) = \mathbf{E}(Y) = 0$ and $\mathbf{E}(Z) = \frac{\rho\sigma y}{\tau}$, the covariance is

$$\begin{aligned}
\mathbf{Cov}(X,Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f(x,y)\,dx \\
&= \frac{1}{\tau\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2\tau^2}} \left(\frac{1}{\sigma\sqrt{2\pi(1-\rho^2)}} \int_{-\infty}^{\infty} x e^{-\frac{1}{2(1-\rho^2)\sigma^2}\left(x - \frac{\rho\sigma y}{\tau}\right)^2}\,dx\right) dy \\
&= \frac{1}{\tau\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\rho\sigma}{\tau} y^2 e^{-\frac{y^2}{2\tau^2}}\,dy = \frac{\rho\sigma}{\tau}\,\mathbf{Var}(Y) = \rho\sigma\tau.
\end{aligned}$$

The correlation coefficient is thus

$$\frac{\mathbf{Cov}(X,Y)}{\sqrt{\mathbf{Var}(X)\ \mathbf{Var}(Y)}} = \frac{\rho\sigma\tau}{\sigma\tau} = \rho.$$