Week 2, Question 4

4. **Gamma and Beta Random Variables**
Let $V_1 \sim \text{Gamma}(a, \lambda)$ be independent of $V_2 \sim \text{Gamma}(b, \lambda)$. These variables may be used to define the Beta, F^* and F distributions.

a) Define $S = V_1 + V_2$ and $X = V_1/(V_1 + V_2)$. Find the joint pdf for S and X and show that these are independent $\text{Gamma}(a + b, \lambda)$ and $\text{Beta}(a, b)$ random variables. Explain what this implies about the waiting time for some number of Poisson events, and the proportion of that time spent waiting for the first event (e.g.). Give the analogous explanation in terms of squared, centered Normal random variables.

Define $S = V_1 + V_2$ and $X = V_1/(V_1+V_2)$

Define $V_1 = X(S)$ and $V_2 = S(1-X)$

Jacobian Transformation: To go from $f_{V_1V_2}(v_1,v_2)\,dv_1dv_2$ to $f_{SX}(s,x)\,ds\,dx$

$$\frac{d(v_1,v_2)}{d(s,x)} = \begin{pmatrix} \frac{dv_1}{ds} & \frac{dv_1}{dx} \\ \frac{dv_2}{ds} & \frac{dv_2}{dx} \end{pmatrix} = \begin{pmatrix} x & s \\ 1-x & -s \end{pmatrix} \therefore |\det| = |-sx - s(1-x)| = |-s| = s$$

So $f_{V_1V_2}(v_1,v_2) = f_{V_1}(v_1)f_{V_2}(v_2) = \frac{\lambda^a}{\Gamma(a)} v_1^{a-1} e^{-\lambda v_1}\left(\frac{\lambda^b}{\Gamma(b)} v_2^{b-1} e^{-\lambda v_2}\right)$

Substitute $V_1 = XS$ and $V_2 = S(1-X)$

$$f_{SX}(s,x) = \frac{\lambda^a}{\Gamma(a)}(xs)^{a-1} e^{-\lambda xs}\left(\frac{\lambda^b}{\Gamma(b)}(s(1-x))^{b-1} e^{-\lambda(1-x)s}\right)$$

$$= \frac{\lambda^{a+b} s^{a+b-2}}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} e^{-\lambda s}$$

Multiply by S - jacobian constant

$$= \frac{\lambda^{a+b} s^{a+b-1}}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} e^{-\lambda s}$$

Multiply by $\frac{\Gamma(a+b)}{\Gamma(a+b)}$:

$$f_{S,X}(s,x) = \underbrace{\left[\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}\right]}_{\text{Beta}(a,b)} \cdot \underbrace{\left[\frac{\lambda^{a+b}}{\Gamma(a+b)} s^{a+b-1} e^{-\lambda s}\right]}_{\text{Gamma}(a+b,\lambda)}$$

Note: Because they are a product of $f_S(s)$ and $f_X(x)$ they are independent.

S = total waiting time for $a+b$ events

X = proportion of total wait time spent on the 1st a events

So,

$V_1 \sim \text{Gama}(a,\lambda) \rightarrow a = \frac{m_1}{2} \rightarrow V_1 \sim \chi^2(m_1)$

$V_2 \sim \text{Gamma}(b,\lambda) \rightarrow b = \frac{m_2}{2} \rightarrow V_2 \sim \chi^2(m_2)$ $\lambda = \frac{1}{2}$

b) Use LOTUS (the law of the unconscious statistician) and integration by recognition (using the fact that pdf's must integrate to 1) show that $E(X) = \frac{a}{a+b}$.

$X \sim Beta(a,b) \therefore f_X(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} \quad 0<x<1$

LOTUS: For a random variable X w/ pdf $f_X(x)$ $E(g(x)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx$

And we know that:

$$\int_0^1 x^{a-1}(1-x)^{b-1} dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad a>0, b>0$$

$$\text{So, } E(x) = \int_0^1 \underset{g(x)}{x} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1} dx$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \underbrace{x^{a+1-1} (1-x)^{b-1}}_{\propto Beta(a+1,b)} dx$$

*Note from presentation 2:

$\Gamma(a+1) = a\Gamma(a)$ and $\Gamma(a+b+1) = (a+b)\Gamma(a+b)$

$$E(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \left(\frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)} \right)$$

$$= \frac{\cancel{\Gamma(a+b)}}{\cancel{\Gamma(a)}\cancel{\Gamma(b)}} \left(\frac{a\cancel{\Gamma(a)}\cancel{\Gamma(b)}}{(a+b)\cancel{\Gamma(a+b)}} \right) = \frac{a}{a+b}$$

c) Show that Uniform$(0, 1)$ is a special case of the Beta distribution. As an example of a Uniform variable, consider the joint pdf of X and Y, the times of the first and second events from a Poisson process ($X < Y$). Show that the conditional distribution of $X|Y = y$ is Uniform by writing $X = V_1$ and $Y = V_1 + V_2$, for V_1 and V_2 iid Exponential variables. Note how $X|Y = y$ is Uniform, but for a constant y, the distribution of $X|X < y$ is *not* Uniform.

$V_1, V_2 \overset{iid}{\sim}$ Exponential$(\lambda) \sim$ Gamma$(1, \lambda)$

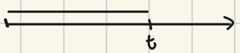
we know from part a that:

$\frac{V_1}{V_1+V_2} \sim$ Beta $(1,1)$ or Uniform $(0,1)$ special case where $a=b=1$

V_1= time until 1^{st} event

V_2= time between 1^{st} and 2^{nd} event

$V_1 | V_1+V_2 = t \sim$ Uniform $(0, t)$



- If we know how long the total events took (V_1+V_2), all we know about event 1 is that it happened at some point before t. But all times are equally likely so $t \sim$ Uniform $(0,t)$

However,

Not the same as $V_1 | V_1 < y$ is NOT Uniform.

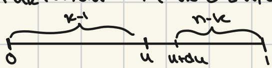
- we do not know when V_1 happened. It is most likely close to time 0 but could happen anytime before t.
- So in this case $V_1 | V_1 < t \sim$ Expo(λ)

d) For $U_1, \ldots, U_n \overset{\text{i.i.d.}}{\sim} \text{Unif}(0,1)$, define $U_{(k)}$ to be the kth order statistic, with $0 < U_{(1)} < \ldots < U_{(n)} < 1$. Use a differential argument to show that $U_{(k)} \sim \text{Beta}(k, n-k+1)$, for $k = 1, \ldots, n$.

$U_1, \ldots, U_n \overset{\text{iid}}{\sim} \text{Unif}(0,1)$

$U_k = k^{th}$ order statistic (ie, $U_1 = \min(U_1, \ldots U_n)$ and $U_n = \max(U_1, \ldots U_n)$)

$f_{U_k}(u)du \approx P(U_k \in [u, u+du])$



$P(U_k \in [u, u+du]) = P(k\text{-}1 \; U_i\text{s} < u, \; 1 \; U_i \in [u, u+du], \; n\text{-}k \; U_i\text{s} > u+du)$

note: any n choices could be the n in $[u, u+du]$

note: $\binom{n-1}{k-1}$ choices for the lower k-1

So, $P(U_k \in [u, u+du]) = n\binom{n-1}{k-1} u^{k-1}(du)\,(1-(u+du))^{n-k}$

$f_{U_k}(u)du \approx P(U_k \in [u, u+du])$

$\lim_{du \to 0}$ and divide du then:

$$f_{U_k}(u) = n\binom{n-1}{k-1} u^{n-1}(1-u)^{n-k} \quad 0<u<1$$

$$= \frac{n(n-1)!}{(k-1)!(n-k-2)!} u^{k-1}(1-u)^{n-k}$$

$$= \frac{n!}{(k-1)!(n-k)!} u^{k-1}(1-u)^{n-k}$$

$$= \frac{\Gamma(k+n-k+1)}{\Gamma(k)\Gamma(n-k+1)} u^{k-1}(1-u)^{n-k}$$

Simplify to mirror Beta:

$$\frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)} = \frac{\Gamma(k+n-k+1)}{\Gamma(k)\Gamma(n-k+1)} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$

↳ Beta distribution where $a = k$ and $b = n-k+1$

So, $U_k \sim \text{Beta}(k, n-k+1)$