Example: Suppose $X, Y \overset{iid}{\sim} Exp(\lambda)$.

(a) Find the distribution of $Z := X/Y$.

(b) Using (a) or otherwise, find the distribution of $U := \dfrac{X}{X+Y}$.

Remark: We have already solved Part (a) by finding the cdf of Z and then finding a pdf of Z; see pg 53, 54, 59, 60 for details. Now we shall use the theorem (or more specifically, the last corollary) stated in Pg 107 and ~~show~~ see how short the calculation becomes. This is because the hard work is hidden in the proof of this theorem.

Solution: Note that $\text{Range}(X) = \text{Range}(Y) = (0,\infty)$ and hence $\text{Range}(Z) \subseteq (0,\infty)$. Also X, Y are iid with each following $Exp(\lambda)$ implies (X,Y) has a joint density function

$$f_{X,Y}(x,y) = \lambda^2 e^{-\lambda(x+y)} \quad \text{if } x>0, y>0.$$

Therefore, using the last corollary of Pg 107, it follows that Z is a cont r.v. with a

112

pdf

$$f_Z(z) = \int_0^\infty y f_{X,Y}(yz, y)\, dy, \quad z > 0.$$

Take $z \in (0,\infty)$. Then we have

$$f_Z(z) = \int_0^\infty y \lambda^2 e^{-\lambda(yz+y)}\, dy$$

$$= \lambda \int_0^\infty y \lambda e^{-\lambda(1+z)y}\, dy$$

$$= \frac{\lambda}{1+z} \int_0^\infty y\, \lambda(1+z)\, e^{-\lambda(1+z)y}\, dy$$

$$= \frac{\lambda}{1+z} E(Y') \qquad \left[\text{Here } Y' \sim \text{Exp}(\lambda(1+z))\right]$$

$$= \frac{\lambda}{1+z} \cdot \frac{1}{\lambda(1+z)} = \frac{1}{(1+z)^2}.$$

Therefore Z is a cont r.v. with a pdf

$$f_Z(z) = \frac{1}{(1+z)^2} \quad \text{if } z > 0.$$

(b) $\text{Range}(X,Y) = (0,\infty)\times(0,\infty)$

$$\Rightarrow \quad U = \frac{X}{X+Y} \quad \text{has} \quad \text{Range}(U) \subseteq (0,1).$$

Also $$U = \frac{X/Y}{1+X/Y} = \frac{Z}{1+Z}.$$

We shall try to use the change of density ~~forula~~ formula + Part (a) to compute a pdf of U.

To this end, take $I = \text{Range}(Z) = (0,\infty)$

and $g : I \to \mathbb{R}$ defined by

$$g(z) = \frac{z}{1+z}, \quad z \in I = (0,\infty).$$

Exc: Check that $\text{Range}(g) = (0,1) =: J$ and $g : I \to J$ is a diffble bijection such that

$$g'(z) = \frac{1}{(1+z)^2} > 0 \quad \forall\, z \in I.$$

Exc: Check that the inverse map of g is given by $g^{-1} : J \to I$

$$g^{-1}(u) = \frac{u}{1-u}, \quad u \in J = (0,1).$$

In particular, $\dfrac{d}{du} g^{-1}(u) = \dfrac{1}{(1-u)^2}, \quad u \in J.$

By the remark given in Pg 72, $\text{Range}(U) = J = (0,1)$.

~~Hence take $u \in (0,1)$.~~ By the change of density formula, (see Pg 72), U is a cont r.v. and a pdf of U is given by

$$f_U(u) = \begin{cases} f_Z\left(g^{-1}(u)\right)\left|\dfrac{dg^{-1}(u)}{du}\right| & \text{if } u \in (0,1), \\ 0 & \text{if } u \notin (0,1). \end{cases}$$

Take $u \in (0,1)$. Then we have

$$f_U(u) = f_Z\left(\frac{u}{1-u}\right) \cdot \frac{1}{(1-u)^2}$$

$$= \frac{1}{\left(1 + \frac{u}{1-u}\right)^2} \cdot \frac{1}{(1-u)^2}$$

$$= \frac{1}{\left(\frac{1}{1-u}\right)^2} \cdot \frac{1}{(1-u)^2} = 1$$

Therefore U is a cont r.v. with a pdf $f_U(u) = 1$ if $u \in (0,1)$. Therefore, $U = X/(X+Y) \sim \text{Unif}(0,1)$.

We have therefore proved the following result.

Propn: If $X, Y \overset{iid}{\sim} \text{Exp}(\lambda)$, then

$$\frac{X}{X+Y} \sim \text{Unif}(0,1).$$

Exc: ~~Show~~ Prove the above propn by first computing the cdf of $U = \frac{X}{X+Y}$ and then finding a pdf of U.

Remarks: ① We shall show later that

$$(X+Y) \perp\!\!\!\perp \frac{X}{X+Y}.$$

This will need the change of joint density formula.

② In fact, we shall show the following ~~fact~~ more general result:

$$\left.\begin{array}{l} X \sim \text{Gamma}(\alpha, \lambda) \\ Y \sim \text{Gamma}(\beta, \lambda) \end{array}\right\} \text{ind} \Rightarrow \text{ind} \left\{\begin{array}{l} X+Y \sim \text{Gamma}(\alpha+\beta, \lambda) \\ \frac{X}{X+Y} \sim \text{Beta}(\alpha, \beta) \end{array}\right.$$

③ The independence of $X+Y$ and $\frac{X}{X+Y}$ also follows from a very deep theorem (known as Basu's Theorem) in theoretical statistics that you would perhaps learn in a statistics course in future.

Exc: If $X, Y \overset{iid}{\sim} N(0,1)$, then find the dist^n of $Z = \frac{X}{Y}$. Does Z have a finite mean? Please justify your answer.

Two more families of distributions important in statistics

1) Student's t-distribution or t-distribution

Def^n: Suppose $Z \sim N(0,1)$, $X \sim \chi^2_n$ and $X \perp\!\!\!\perp Z$. Then the r.v.

$$T := \frac{Z}{+\sqrt{X/n}}$$

is said to follow t-distribution with n degrees of freedom.

Notation: In the above situation, we write

$T \sim t_n$.

Remarks: ① Note that Student's t-distribution borrows its degree of freedom from the χ^2-distⁿ used in its definition.

② It is possible to compute a pdf for the t-distⁿ with n degrees of freedom but we shall skip it because nobody uses its pdf.

③ Suppose $X_1, X_2, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$.

Then the following results can be shown to hold:

ⓐ $$Z := \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i - \mu}{\sigma/\sqrt{n}} \sim N(0,1).$$

ⓑ $$V := \frac{(n-1)S^2}{\sigma^2} := \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2} \sim \chi^2_{n-1}.$$

ⓒ $$\overline{X} \perp\!\!\!\perp S^2 \Rightarrow Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \perp\!\!\!\perp \frac{(n-1)S^2}{\sigma^2} = V$$

d From a, b, c above, we get

$$\text{ind} \begin{cases} Z = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1) \\ V = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1} \end{cases}$$

$$\Rightarrow T := \frac{Z}{+\sqrt{V/(n-1)}} = \frac{\sqrt{n}(\bar{X}-\mu)/\sigma}{+\sqrt{S^2/\sigma^2}}$$

$$= \frac{\sqrt{n}(\bar{X}-\mu)}{S}$$

$$= \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1},$$

where $S = +\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$ = Sample Standard Deviation.

The fact $T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ is used "to perform statistical inference" on the unknown parameter μ ~~from~~ using a random (i.e., iid) sample ~~$X_1, X_2, \ldots, X_n$~~

$X_1, X_2, \ldots, X_n$ from $N(\mu, \sigma^2)$ dist^n when σ^2 is also unknown.

2) F-distribution or Snedecor's F-distribution or Fisher-Snedecor distribution

Defn: Suppose $S_1 \sim \chi^2_{d_1}$, ~~and~~ $S_2 \sim \chi^2_{d_2}$ and $S_1 \perp\!\!\!\perp S_2$. The the r.v.

$$W := \frac{S_1/d_1}{S_2/d_2}.$$

is said to follow F-distribution with d_1 and d_2 degrees of freedom.

Notation: $W \sim F_{d_1, d_2}$.

Remarks: ① F-dist^n also borrows its degrees of freedom from the underlying chi-squared r.v.s.

② A pdf of F-distⁿ ~~can~~ with d_1 and d_2 degrees of freedom can also be computed but nobody will use it. Hence we shall skip this computation as well.

③ F-distⁿ arises naturally in an area of statistics called Analysis of Variance (ANOVA). You will learn in details about this topic in a course in statistics.

Exc: Suppose (X, Y) is uniformly distributed on the unit disk $D = \{(x,y) \in \mathbb{R}^2 : x^2+y^2<1\}$. Let $S = X^2+Y^2$. Find the distⁿ of $W = -2 \log_e S$.

[Hint: First find the cdf of S]

Exc: Suppose $U \sim \text{Unif}(0,1)$. Find the distⁿ of the following r.v.s.

(i) $X = U^2$ (ii) $Y = +\sqrt{U}$ (iii) $Z = \frac{1}{U}$.