Class Notes 3-11-2019

- **Linear Transformations** $g(X) = aX + b$**,** $a \neq 0$**:**
Let X be a continuous random variable with cdf $F_X(x)$ and pdf $f_X(x)$. In Problem 4 of HW 5 you worked out the cdf and pdf of $Y = g(X) = bX + a$. For $b > 0$ we have

$$F_Y(y) = P(Y \leq y) = P(bX + a \leq y) = P\left(X \leq \frac{y-a}{b}\right) = F_X\left(\frac{y-a}{b}\right)$$

with pdf $f_Y(y) = \frac{1}{b} f_X\left(\frac{y-a}{b}\right)$. When $b < 0$ we get

$$F_Y(y) = P(Y \leq y) = P(bX + a \leq y) = P\left(X \geq \frac{y-a}{b}\right) = 1 - F_X\left(\frac{y-a}{b}\right)$$

with pdf $f_Y(y) = -\frac{1}{b} f_X\left(\frac{y-a}{b}\right) = \frac{1}{|b|} f_X\left(\frac{y-a}{b}\right)$ where the latter form also holds for $b > 0$.

Comment: Such densities of the form $f_Y(y) = \frac{1}{|b|} f_X\left(\frac{y-a}{b}\right)$ with f_X some known standard density such as $f_X = \varphi(x)$ (standard normal density) are very useful models when $Y_1, \ldots, Y_n$ iid $\sim Y$ are observed but the transformation parameters a and b in the linear transform $Y_i = a + bX_i$ are unknown. If we denote $\bar{Y} = \frac{1}{n}\sum_{i=1}^n Y_i = a + b\bar{X}$ we have

$$\frac{Y_i - \bar{Y}}{b} = \frac{a + bX_i - (a + b\bar{X})}{b} = \frac{bX_i - b\bar{X}}{b} = X_i - \bar{X} \qquad \text{and} \qquad \frac{\bar{Y} - a}{b} = \bar{X}$$

and thus

$$\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{b^2} = \sum_{i=1}^n (X_i - \bar{X})^2$$

and since the X_i have a known distribution the distribution of $\sum_{i=1}^n (X_i - \bar{X})^2$ can either be derived mathematically or to arbitrary accuracy by simulating lots of samples $X_1, \ldots, X_n \sim f_X(x)$ and computing $\sum_{i=1}^n (X_i - \bar{X})^2$ each time. From such a distribution (exact or simulated) we could then get values D and G such that

$$.95 = (\approx) P\left(D \leq \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{b^2} \leq G\right) = P\left(\frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{G} \leq b^2 \leq \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{D}\right)$$

i.e., the interval $[\sum_{i=1}^n (Y_i - \bar{Y})^2/G, \sum_{i=1}^n (Y_i - \bar{Y})^2/D]$ captures the unknown parameter b^2 with probability .95 and by taking square roots the parameter b with probability .95. Such intervals are referred to as *confidence intervals* for b^2 or b. It is the intervals whose randomness is expressed by the probability statement. The parameters $b^2(b)$ are not random and simply unknown. In any specific instance we will not know whether our interval captured the unknown parameter or not. We only have the assurance (confidence) that in 95% of the cases of applying this method we would have captured our target. In this context it is worth quoting Myles Hollander: "Statistics means never having to say you are certain,"

In a similar fashion you can exploit the known or simulated distribution of

$$\frac{\bar{Y} - a}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} = \frac{(\bar{Y} - a)/b}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2/b^2}} = \frac{\bar{X}}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}}$$

to get

$$
\begin{aligned}
.95 &= (\approx)P\left(D \le \frac{\bar{Y}-a}{\sqrt{\sum_{i=1}^n (Y_i-\bar{Y})^2}} \le G\right) \\
&= (\approx)P\left(\bar{Y} - G\sqrt{\sum_{i=1}^n (Y_i-\bar{Y})^2} \le a \le \bar{Y} - D\sqrt{\sum_{i=1}^n (Y_i-\bar{Y})^2}\right)
\end{aligned}
$$

for a confidence interval for a.

- **A Monotone Transformation of a Continuous Random Variable:**
  Let $U \sim \text{Unif}[0,1]$ and let $Y = -\ln(1-U)$. Since $1-U \sim \text{Unif}[0,1]$ as well, we have that $Y = -\ln(1-U)$ and $Y^* = -\ln(U)$ have the same distribution, namely for $y > 0$

  $$P(-\ln(U) > y) = P(U < e^{-y}) = e^{-y} \quad \text{so that} \quad F_{Y^*}(y) = P(-\ln(U) \le y) = 1 - e^{-y} \quad \text{for } y > 0$$

  while $F_{Y^*}(y) = 0$ for $y \le 0$. Hence both $Y = -\ln(1-U)$ and $Y^* = -\ln(U)$ have an exponential distribution with mean 1.

- **A Non-monotone Transform of $Z \sim \mathcal{N}(0,1)$: $Y = g(Z) = Z^2$**

  $$F_Y(y) = P(Y \le y) = P(Z^2 \le y) = P(-\sqrt{y} \le Z \le \sqrt{y}) = 2\Phi(\sqrt{y}) - 1 \ \text{ for } y > 0$$

  and $F_Y(y) = 0$ for $y \le 0$ and with density

  $$f_Y(y) = \frac{1}{\sqrt{y}}\varphi(\sqrt{y})I_{(0,\infty)}(y) = \frac{1}{\sqrt{2\pi y}}e^{-y/2}I_{(0,\infty)}(y)$$

  The distribution of Z^2 is said to have a chi-squared distribution with one degree of freedom. We also write $Z^2 \sim \chi_1^2$.

- **General Monotone Transformations:**
  Let X be a continuous random variable with pdf $f_X(x)$. Let g be a monotone increasing function with derivative $\neq 0$ except at finitely many points so that it has an inverse $g^{-1}(y)$. Then $Y = g(X)$ has cdf

  $$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \le g^{-1}(y)) = F_X(g^{-1}(y))$$

  for y in the range $D = (A, B)$ of $g(x)$. Here A and/or B may be $\pm\infty$, respectively. Below that range $F_Y(y) = 0$ and above that range $F_Y(y) = 1$. By differentiation we get the corresponding density as

  $$f_Y(y) = f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y)I_{(A,B)}(y) = f_X(g^{-1}(y))\frac{1}{g'(g^{-1}(y))}I_{(A,B)}(y)$$

  with zero density outside (A, B). The derivative of the inverse $g^{-1}(y)$ follows from

  $$y = g\left(g^{-1}(y)\right) \qquad \Longrightarrow \qquad \frac{d}{dy}g\left(g^{-1}(y)\right) = 1 = g'(g^{-1}(y))\frac{d}{dy}g^{-1}(y)$$

  When g be a monotone decreasing function with derivative $\neq 0$ except at finitely many points so that it has an inverse $g^{-1}(y)$. Then $Y = g(X)$ has cdf

  $$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P(X \ge g^{-1}(y)) = 1 - F_X(g^{-1}(y))$$

for y in the range $D = (A, B)$ of $g(x)$. Here A and/or B may be $\pm\infty$, respectively. Below that range $F_Y(y) = 0$ and above that range $F_Y(y) = 1$. By differentiation we get the corresponding density as

$$f_Y(y) = -f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y)I_{(A,B)}(y) = -f_X(g^{-1}(y))\frac{1}{g'(g^{-1}(y))}I_{(A,B)}(y)$$

with zero density outside (A, B). The two cases can be combined in the following way for any monotone differentiable function $Y = g(X)$ with zero derivative at at most a finite number of places.

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right| I_{(A,B)}(y) = f_X(g^{-1}(y))\frac{1}{|g'(g^{-1}(y))|}I_{(A,B)}(y)$$

with zero density outside (A, B).

- **Weibull-Gumbel Example:**
  Let $X \sim \text{Weib}(\alpha, \beta)$ for $\alpha > 0$ abd $\beta > 0$, i.e., $F_X(x) = 1 - e^{-(x/\alpha)^\beta}$ for $x \geq 0$. Find cdf and density of $Y = \ln(X) = g(X)$. By differentiation the Weibull density is

$$f_X(x) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(x/\alpha)^\beta} I_{[0,\infty)}(x)$$

  We have $g^{-1}(y) = e^y$ with derivative e^y, so that for all $y \in \mathbb{R}$

$$f_Y(y) = \frac{\beta}{\alpha}\left(\frac{e^y}{\alpha}\right)^{\beta-1} e^{-(e^y/\alpha)^\beta} I_{[0,\infty)}(e^y)e^y = \beta e^{-(e^y/\alpha)^\beta}(e^y/\alpha)^\beta = \frac{1}{b}e^{\frac{y-a}{b}}e^{-e^{\frac{y-a}{b}}} = \frac{1}{b}g\left(\frac{y-a}{b}\right)$$

  where $a = \log\alpha$ and $b = 1/\beta$ and

$$g(x) = e^{-e^x}e^x \qquad \text{is the standard Gumbel density}$$

  and the parametrization with (a, b) is much easier to understand in its location/scale character.

  A more direct way is to derive the cdf of Y as follows

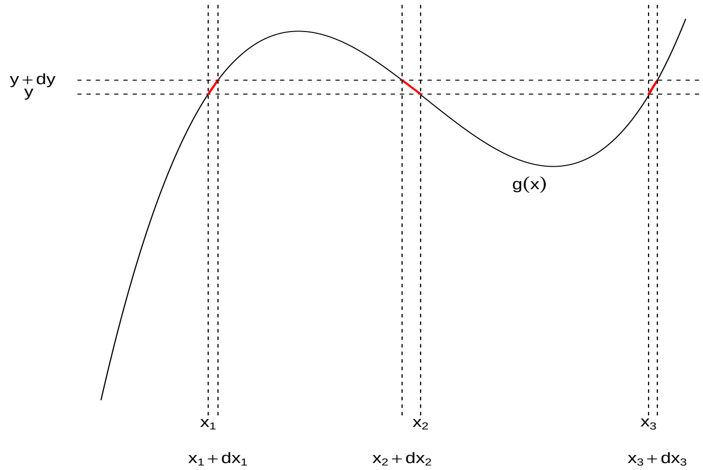
$$P(Y > y) = P\left(X > e^y\right) = e^{-\left(\frac{e^y}{\alpha}\right)^\beta} = e^{-e^{\frac{y-\log\alpha}{1/\beta}}} = e^{-e^{\frac{y-a}{b}}} \quad \text{with} \quad F_Y(y) = P(Y \leq y) = 1 - e^{-e^{\frac{y-a}{b}}}$$

  and the density is obtained by differentiation.

- **General Formula for Density of $Y = g(X)$ when g is not 1-1**

$$f_Y(y) = \sum_{\substack{x:g(x)=y \\ g'(x)\neq 0}} f_X(x)\frac{1}{|g'(x)|}$$

  Here we do not replace x in $f_X(x)$ and $g'(x)$ by $x = g^{-1}(y)$ because g does not necessarily have an inverse. There may be several values x that map into the same y. That is the reason for the notation used.
  The above formula for $f_Y(y)$ can easily be derived via the infinitesimal interpretation of $f_Y(y)dy$ as approximation to $P(Y \in [y, y + dy])$, i.e., $P(Y \in [y, y + dy]) \approx f_Y(y)dy$



In the plot the x_i abscissa values correspond to the intersect locations of the non-monotone function with the horizontal y line, i.e., $g(x_i) = y$.