4.1 We know that $\mu_w = 150$, $\sigma_w = 30$, $r = 0.9$
$\mu_a = 0.8$, $\sigma_a = 0.1$

a. Given $w_0 = 170$, find a_0 (assuming $\hat{a} = r\hat{w}$)

$$\frac{a^p - \mu_a}{\sigma_a} = \frac{r(w_0 - \mu_w)}{\sigma_w}$$

$$\Rightarrow a^p = \frac{r\sigma_a(w_0 - \mu_w)}{\sigma_w} + \mu_a$$

So, $a_0^p \approx 0.86$

b. Using the same procedure as above, but with $\hat{w} = r\hat{a}$, we find that $w_0^p \approx 136.5$

c. I expect it to be very reliable as the correlation coefficient is large at 0.9, which is greater than the 0.5 standard.

4.2 We know that $\mu_F = \$60{,}000$, $\sigma_F = \$20{,}000$
$\mu_I = 100$, $\sigma_I = 15$, $r = 0.3$

a. Using $I_0 = \$70{,}000$, find Q_0^p (assuming that $\hat{F} = r\hat{I}$)

$$\frac{F_0^p - \mu_F}{\sigma_F} = r \times \frac{I_0 - \mu_I}{\sigma_I}, \text{ so } Q_0^p \approx 102.25$$

(continued)

4.2 b. Fairly unreliable, since the correlation coefficient is 0.3, a relatively low value when compared to the standard 0.5.

c. Assuming that the rise is speaking about a new, distinct family in the same sample set with a relatively larger family income ~~that~~ than a previous family, then yes the correlation is positive so a higher family income would indicate a higher IQ prediction. But it is important to ~~[illegible]~~ remember that correlation doesn't mean causation.

4.5 **Objective:** Find a and b

using: $\hat{x}^p = a\hat{y} + b$

$u = \hat{x} - \hat{x}^p$ (error-per-sample)

Assume: $\text{mean}(\{u\}) = 0$ and to minimize $\text{var}(\{u\})$

a. Using the mean of the error terms, find b.

$$\text{mean}(\{u\}) = 0 \Rightarrow \text{mean}(\{\hat{x} - \hat{x}^p\}) = 0$$

We know that $\text{mean}(\{\hat{x}\}) = 0$, as is $\text{mean}(\{\hat{y}\})$

$$\Rightarrow \text{mean}(\{\hat{x}\}) - \text{mean}(\{\hat{x}^p\})$$

$$\Rightarrow -\text{mean}(\{\hat{x}^p\}) = -\text{mean}(\{a\hat{y} + b\})$$

$$= -a\,\text{mean}(\{\hat{y}\}) - b$$

So, $b = 0$

b. We want to minimize the variance to find a.

$$\text{var}(\{u\}) = \text{mean}(\{(u - \text{mean}(\{u\}))^2\})$$

$$= \text{mean}(\{u^2\})$$

$$= \text{mean}(\{(\hat{x} - \hat{x}^p)\})$$

$$= \text{mean}(\{(\hat{x} - (a\hat{y}))^2\})$$

$$= \text{mean}(\{\hat{x}^2 - 2a\hat{x}\hat{y} + a^2\hat{y}^2\})$$

$$= \underbrace{\text{mean}(\{\hat{x}^2\})}_{1} - 2a\,\underbrace{\text{mean}(\{\hat{x}\hat{y}\})}_{r} + a^2\,\underbrace{\text{mean}(\{\hat{y}^2\})}_{1}$$

$$= 1 - 2a \cdot r + a^2 \cdot 1$$

$\text{var}(\{u\}) \Rightarrow 1 - 2ar + a^2$, so $\dfrac{\partial\,\text{var}(\{u\})}{\partial a} = -2r + 2a = 0$

So, $a = r$

4.6 We know that $\mu_Y = 1988.5$, $\sigma_Y = 14$, $r = 0.882$

$\mu_T = 0.175$, $\sigma_T = 0.231$,

(Y = year, T = Earth Temp.)

$T_0^P(y = 2014) \approx 0.550$

$T_1^P(y = 2028) \approx 0.756$

$T_2^P(y = 2042) \approx 0.962$

4.7 We know that $\mu_T = 0.175$, $\sigma_T = 0.231$

$\mu_N = 31.6$, $\sigma_N = 30.8$, $r = 0.471$

(N = Number of tornadoes, T = Earth temp.)

$N_0^P(T = 0.5) \approx 51.21$

$N_1^P(T = 0.6) \approx 57.49$

$N_2^P(T = 0.7) \approx 63.77$