22.1. a. $Y = x'\theta + e, \quad e \sim f(e)$

$$
\begin{aligned}
Pr(y_i \le y \mid x_i = x) &= Pr(x_i'\theta + e_i \le y \mid x_i = x) \\
&= Pr(e_i \le y - x_i'\theta \mid x_i = x) \\
&= Pr(e_i \le y - x_i'\theta) \quad \text{since } e, X \text{ ind.} \\
&= F(y - x_i'\theta)
\end{aligned}
$$

By differentiation, $f(y \mid x) = f(y - x'\theta)$

b. $p_i = -\log f(y_i - x_i\theta)$

$$
\begin{aligned}
\psi_i &= \partial/\partial\theta \; p_i \\
&= \frac{f'(y_i - x_i\theta)}{f(y_i - x_i\theta)} x_i
\end{aligned}
$$

c.
$$
\begin{aligned}
\Omega &= E[\psi_i \psi_i'] \\
&= E\left[\left(\frac{f'(y_i - x_i\theta)}{f(y_i - x_i\theta)}\right)^2 x_i x_i'\right] \\
&= E\left[\left(\frac{f'(e_i)}{f(e_i)}\right)^2 x_i x_i'\right] \\
&= E\left[\left(\frac{f'(e_i)}{f(e_i)}\right)^2\right] E[x_i x_i']
\end{aligned}
$$

$V = \Omega^{-1}$

23.1. a. $Y = \exp(\theta) + e, \quad E[e] = 0$

Since $\exp(\theta)$ is nonlinear, the conditional mean is nonlinear in θ. This is a nonlinear regression model.

b. Let $\gamma = \exp(\theta)$ and consider $Y = \gamma + e$.
We can estimate $\hat{\gamma} = \bar{Y}_n$.
$\hat{\theta} = \log(\hat{\gamma})$.

c. This is the same as the NLLS estimator since M-estimators don't change when reparameterized.

23.2 $\quad y^{(\lambda)} = \beta_0 + \beta_1 x + e, \quad E[e|x] = 0$

This is not a linear model.

$$y(\lambda) = \begin{cases} \dfrac{y^{\lambda} - 1}{\lambda} & \lambda > 0 \\ \log y & \lambda = 0 \end{cases}$$

23.7. $\quad Y = m(X, \theta) + e \qquad E[e|X] = 0$

The SE is: $s = \sqrt{\dfrac{\partial m(x,\theta)}{\partial \theta} \hat{V} \dfrac{\partial m(x,\theta)}{\partial \theta'}}$

The 95% CI is:

$$CI = [m(x,\theta) - 1.96s, \; m(x,\theta) + 1.96s]$$

23.8.

Source	SS	df	MS
Model	768.35943	3	256.119809
Residual	12.713675	334	.038064896
Total	781.0731	337	2.31772434

Number of obs	=	338
R-squared	=	0.9837
Adj R-squared	=	0.9836
Root MSE	=	.1951023
Res. dev.	=	-149.5618

lny	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
/beta	-12.58216	.1845737	-68.17	0.000	-12.94523	-12.21909
/nu	1.042579	.007826	133.22	0.000	1.027185	1.057974
/rho	.4114824	.0584956	7.03	0.000	.2964163	.5265486
/alpha	.3194286	.0115813	27.58	0.000	.2966471	.3422102

$$\hat{\sigma} = \frac{1}{1-\hat{\rho}} = \frac{1}{1-.411} \approx 1.7$$

Other than $\hat{\beta}$, the estimated coefficients are really close.

24.3. $\psi(x) = \tau - 1\{x<0\}$, $E[\psi(Y-\theta)] = 0$

$$E[\psi(Y-\theta)] = E[\tau - 1\{Y-\theta<0\}]$$
$$= \tau - E[1\{Y-\theta<0\}]$$
$$= \tau - E[1\{Y<\theta\}]$$
$$= \tau - P(Y<\theta)$$
$$= 0$$

$\rightarrow$ $P(Y-\theta) = \tau$, so θ is the τ quantile of Y.

24.4. a. $Y = X'\beta + e$, $E[e|X] = 0$

$$E[Y|X] = E[X'\beta + e|X]$$
$$= E[X'\beta|X] + E[e|X]$$
$$= X'\beta + 0$$
$$= X'\beta$$

$\text{med}[Y|X] = X'\beta$ since e is symmetric about 0.

b. The true coefficient β will be the same in both regressions, but the sample estimation $\hat{\beta}$ may differ since the models are minimizing different criteria.

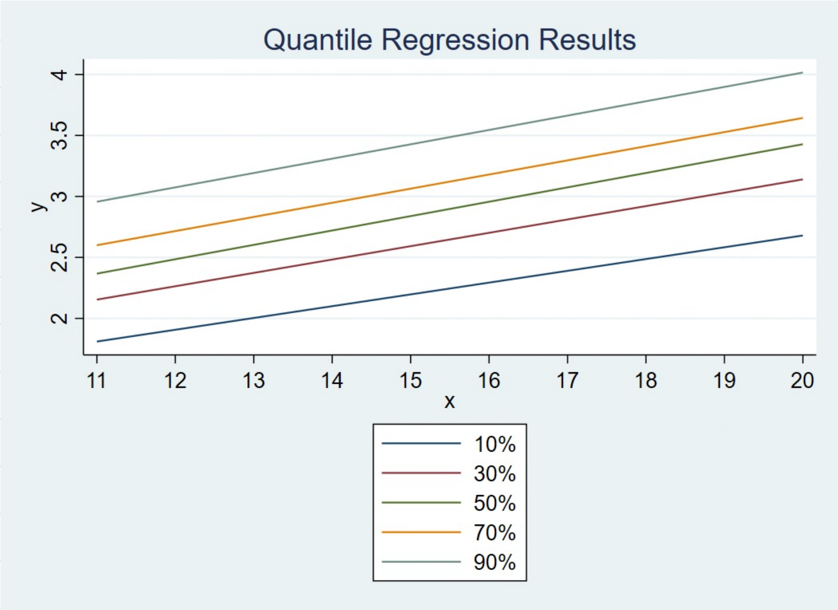
c. We would prefer LAD if we're concerned about outliers.

We would prefer OLS otherwise since it is conditionally unbiased, efficient, and easier computationally.

24.5. $Y = X'\beta + e$

No, the R^2 for $\hat{\beta}_{OLS}$ is going to be higher than R^2 for $\hat{\beta}_{LAD}$ by construction. This is not a good test for determining which model to use.

24.14



The graph above shows quantiles of log wages based on education. The results show that the distribution is approximately homoskedastic.