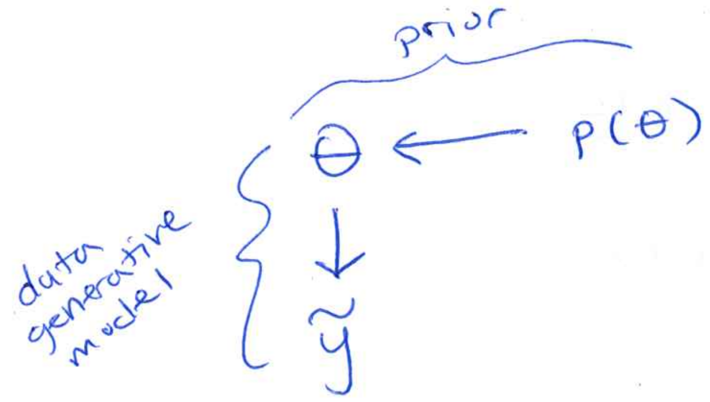
PRIOR PREDICTIVE DISTRIBUTION

- tells us how prior beliefs about θ, $p(\theta)$, translate into beliefs about $\tilde{y}$, according to our sampling model.

($\tilde{y}$ ← future data)

PICTURE:



MATH:

$$p(\tilde{y}) = \int p(\tilde{y} \mid \theta)\, p(\theta)\, d\theta$$

$p(\tilde{y})$: prior predictive distribution

$p(\tilde{y} \mid \theta)\, p(\theta)$ ✓ ✓ $= p(\tilde{y}, \theta)$

pseudo-code:

```
Let S be large #
for (s in 1:S) {
    draw θ^(s) ~ p(θ)
    draw ỹ^(s) ~ p(ỹ | θ^(s))
}
```

To approx $p(\tilde{y})$

POSTERIOR PREDICTIVE DISTR.

- tells us how our <u>posterior</u> beliefs about θ, $p(\theta \mid \underline{y})$, translate into beliefs about $\tilde{y}$.
- useful for evaluating ~~the~~ our sampling model, $p(y \mid \theta)$
- Does not "rule in" our model but may expose a flawed model. Note: this depends on many factors e.g. statistic of interest.

<u>PICTURE</u>:

$y_1, \dots, y_n \equiv \underline{y}$

$\downarrow$

$p(\theta) \searrow \theta$

$\downarrow$

$\tilde{y}$.

<u>MATH</u>:

$$p(\tilde{y} \mid y_1, \dots, y_n) = \int p(\tilde{y}, \theta \mid \underline{y})\, d\theta$$

$$= \int p(\tilde{y} \mid \theta, \underline{y})\, p(\theta \mid \underline{y})\, d\theta$$

$$= \int p(\tilde{y} \mid \theta)\, p(\theta \mid \underline{y})\, d\theta$$

<u>Pseudo-code</u>: To approx $p(\tilde{y} \mid \underline{y})$ w/ Monte Carlo:

Let S be large.

for (s in 1:S){ $\theta^{(s)} \sim p(\theta \mid \underline{y})$

$\tilde{y}^{(s)} \sim p(\tilde{y} \mid \theta^{(s)})$

}

Posterior Predictive Checks:

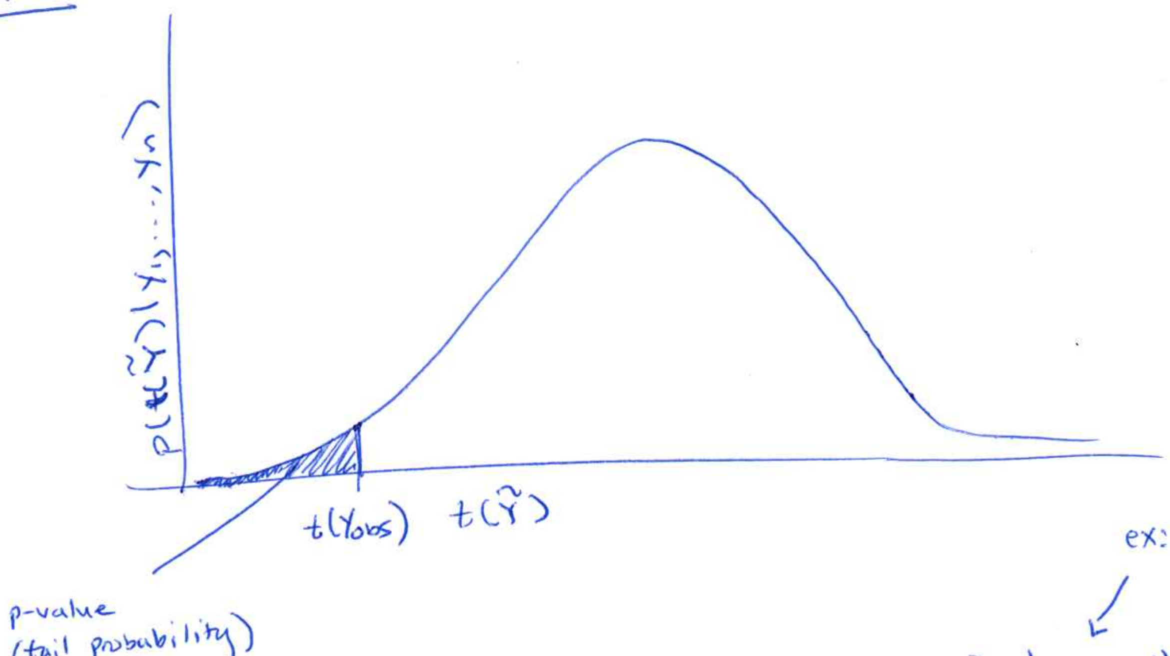
- Given the posterior predictive distr., compute the tail prob of some statistic.

Recall: definition of statistic $t(Y)$ (← random): a function of the data

ex: $t(Y) = \frac{1}{n}\sum Y_i$

ex: $t(Y) = \max(Y_1, \ldots, Y_n)$

Picture:



More generally: p-value: $p = \text{prob}(t(Y) \le t(Y_{obs}) \mid \text{something})$

- something ← ex: $H_0: \theta = \theta_0$ (null Hypothesis)
- something ← $\theta = \hat{\theta}_{MLE}$ = "parametric bootstrap"
- something ← $\{y_1, \ldots, y_n\}$ = "posterior predictive"

Posterior predictive p-value* $= \int_{-\infty}^{t(Y_{obs})} p(t(\tilde{y}) \mid y)\, dt(\tilde{y})$

where $p(t(\tilde{y}) \mid y) = \int_\theta p(t(\tilde{y} \mid \theta)\, p(\theta \mid y)\, d\theta$

* **NOT** distributed unif(0,1) like a traditional p-value

POSTERIOR PREDICTIVE P-VALUE PSEUDOCODE:

To APPROX $prob(t(\tilde{y}) \leq t(y_{obs}) | y)$ using Monte Carlo:

let S be large

for $(s \text{ in } 1:S)$ {

sample $\theta^{(s)} \sim p(\theta | y)$

sample $n = 111$ (here) $\tilde{y} \sim p(\tilde{y} | \theta^{(s)})$

compute & save $t(\tilde{y}_1^{(s)}, \ldots, \tilde{y}_n^{(s)})$

}

Report $mean(t(\tilde{y}) < t(y_{obs}))$