ex) $X \sim \text{Weibull}(\theta, \tau)$

El rev

$$VaR_p(X) = \pi_p = \theta\left[-\ln(1-p)\right]^{\frac{1}{\tau}}$$

Suppose you are given 2 percentiles

$$\pi_p \overset{set}{=} \hat{\pi}_p \text{ and } \pi_q = \theta\left[-\ln(1-q)\right]^{\frac{1}{\tau}} \overset{set}{=} \hat{\pi}_q$$

$$\left(\frac{\pi_p}{\theta}\right)^{\tau} = -\ln(1-p)$$

$$\left(\frac{\pi_q}{\theta}\right)^{\tau} = -\ln(1-q).$$

$$\text{so } \left(\frac{\pi_p}{\pi_q}\right)^{\tau} = \frac{\ln(1-p)}{\ln(1-q)}$$

take ln of both sides and solve for τ

$$\text{so } \hat{\tau} = \frac{\ln\left[\ln(1-p)/\ln(1-q)\right]}{\ln\left(\hat{\pi}_p/\hat{\pi}_q\right)}$$

$$\hat{\theta} = \frac{\hat{\pi}_p}{\left[-\ln(1-p)\right]^{1/\hat{\tau}}}$$

ex) $X \sim LN(\mu, \sigma)$ $\quad \pi_p = e^{\mu + z_p \sigma} = VaR_p(X)$ 14.5

$P(Z \le z_p) = p$

backwards calculation from standard normal table

$z_p = \Phi^{-1}(p)$

$$\ln(\pi_p) = \mu + z_p \sigma \overset{set}{=} \ln(\hat{\pi}_p)$$

$$\ln(\pi_q) = \mu + z_q \sigma \overset{set}{=} \ln(\hat{\pi}_q)$$

$$\ln(\pi_p) - \ln(\pi_q) = \sigma(z_p - z_q) \overset{set}{=} \ln(\hat{\pi}_p) - \ln(\hat{\pi}_q)$$

$$\hat{\sigma} = \frac{\ln(\hat{\pi}_p) - \ln(\hat{\pi}_q)}{z_p - z_q}$$

$$\hat{\mu} = \ln(\hat{\pi}_p) - z_p \hat{\sigma}$$

ex) $X \sim$ inverse exponential (θ)

$$\pi_p = \theta\left[-\ln(p)\right]^{-1} \overset{set}{=} \hat{\pi}_p$$

$$\hat{\theta} = -\hat{\pi}_p \ln(p)$$

See [illegible]
Inverse weibull

17) Sometimes for 2 parameter distributions, one parameter γ_1 is known and the 2nd parameter γ_2 is estimated by matching with π_p.

18) Suppose you have right censored data

$x_1, x_2, \ldots, x_m$, $n-m$ cases censored at $u > x_{(m)}$

order statistics of censored data $x_{(1)}, x_{(2)}, \ldots, x_{(m)}, \underbrace{u, \ldots, u}_{n-m \text{ cases}}$

If $j+1 \le m$, then percentile matching can still be used with $\hat{\pi}_p$ from 15),

ex) For insurance with policy limit $u = 10000$, the ordered observed data are 1000, 1700, 2000, 3500, 4200, 5000, 7300, 4 at the limit

Fit an inverse exponential distribution to the data matching the median, and estimate θ.

Soln: Median $= \pi_{0.5}$ and

or E' rev ex 9)

$\hat{\theta} = -\hat{\pi}_{0.5} \ln(0.5)$ by a previous ex.

$n = 11$, $(n+1)p = 6$

$\hat{\pi}_{0.5} = X_{(6)} = 5000$

so $\hat{\theta} = -5000 \ln(0.5) = \boxed{3465.7359}$

19} If X is (left) truncated at d,

$W = X \mid X > d$ has survival function

$$S_W(x) = \frac{S_X(x)}{S_X(d)} \quad \text{for } X > d.$$

So $S_W(\underbrace{\pi_p(W)}_{VaR_p(W)}) = \dfrac{S_X(\pi_p(W))}{S_X(d)} = 1-p$

or $F_W(\pi_p(W)) = 1 - \dfrac{S_X(\pi_p(W))}{S_X(d)} = p.$

ex} For insurance with a deductible 500, you observe the following losses (including the deductible) (policyholders losses)

600, 1000, 1200, 2000, 3000

Fit a Pareto$(\alpha = 1, \theta)$ dist to the data matching at the median M404 16

and estimate θ. data are from $W = X \mid X > d$

Soln] $\hat{\pi} = \hat{\pi}_{.5}(w) = 1200$ since $(n+1)p = 6(.5) = 3$

$$\frac{S_X(\hat{\pi})}{S_X(500)} \overset{\text{set}}{=} .5 \quad \text{or}$$

[illegible]

$$\frac{\left(\frac{\theta}{\hat{\pi}+\theta}\right)}{\left(\frac{\theta}{500+\theta}\right)} = \frac{500+\theta}{\hat{\pi}+\theta} = 0.5$$

So $500 + \theta = 0.5\,\hat{\pi} + 0.5\,\theta$

$$0.5\,\theta = 0.5\,\hat{\pi} - 500$$

$$\hat{\theta} = \hat{\pi} - 1000 = \boxed{200}$$

20] For a 2 parameter dist X,

could set $\mu_1' = \bar{X}$ and $\pi_p = \hat{\pi}_p$

and solve for γ_1 and γ_2.

21] There are a lot of MLE problems including i) write the likelihood function, ii) write the loglikelihood function, iii) find the MLE $\hat{\theta}$, iv) find an MLE CI for θ, find $\hat{V}(\hat{\theta}) = AV(\hat{\theta})$, perhaps with the delta method.

For these problems, you can have $\underbrace{\text{iid data}}_{}$, grouped data, $\underbrace{\text{censored data}}_{\text{usually right}}$, and $\underbrace{\text{truncated data}}_{\text{usually left}}$.

M483, 580 individual data

22) *For Math 404 (<u>not</u> Math 483, 580)

assume the MLE is a solution

to $\frac{\partial}{\partial \theta_i} \ln L(\underline{\theta}) \stackrel{set}{=} 0 \qquad i = 1, \ldots, k.$

Don't use second derivatives to prove the MLE is the global max since professional actuarial exams are multiple choice.
FAM-S ASTAM formerly C

23) If there is one parameter θ, often the MLE $\hat{\theta} = \overline{w} = \frac{1}{n}\sum_{i=1}^{n} w_i$ where $w_i = t(x_i)$ for some function $t(\cdot)$.

24] Know Central Limit Theorem CLT

Let $X_1, \ldots, X_n$ be iid with $E(X) = \mu$ and $V(X) = \sigma^2$, then

$$\sqrt{n}\,(\bar{X}_n - \mu) \xrightarrow{D} N(0, \sigma^2) \text{ and}$$

$$\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) = \sqrt{n}\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{n\sigma}\right) \xrightarrow{D} N(0,1).$$

25] * The notation $Y_n \xrightarrow{D} X$ means that for large n, we can approximate the cdf of Y_n by the cdf of X. The distribution of X is the limiting distribution or asymptotic distribution of Y_n, and the limiting distribution does not depend on n.

26] * The notation $Y_n \sim AN(\theta, \frac{\tau^2}{n})$, also written $Y_n \approx N(\theta, \frac{\tau^2}{n})$, means approximate the cdf of Y_n as if $Y_n \sim N(\theta, \frac{\tau^2}{n})$.

By the CLT, $\bar{X}_n \sim AN(\mu, \frac{\sigma^2}{n})$. Note that the approximate distribution, unlike the limiting distribution, does depend on n.

27] If the statistic $T_n \sim AN(\gamma, \psi^2)$, (17.5)

then $\pi_p(T_n) = VaR_p(T_n) \approx \gamma + \psi z_p$

where $P(Z \le z_p) = p$ if $Z \sim N(0,1)$.

If $T_n = \bar{X}_n$, $\gamma = \mu$ and $\psi = \frac{\sigma}{\sqrt{n}}$.

28]* For individual data, $x_1, \dots, x_n$ are iid with pdf $f(x)$ or pmf $p(x)$.

So the likelihood function

$$L(\underline{\theta}) = \prod_{i=1}^{n} f(x_i) \quad \text{or} \quad L(\underline{\theta}) = \prod_{i=1}^{n} p(x_i).$$

The log likelihood function

$$\ln L(\underline{\theta}) = \sum_{i=1}^{n} \ln f(x_i) \quad \text{or} \quad \ln L(\underline{\theta}) = \sum_{i=1}^{n} \ln(p(x_i)).$$

29] If $x_1 = 6$, $x_2 = 5$, $x_3 = 7$

and $f(x) = \lambda e^{-\lambda x}$, $= \lambda^n e^{-\lambda \sum x_i}$

then $L(\lambda) = \lambda e^{-\lambda 6} \lambda e^{-\lambda 5} \lambda e^{-\lambda 7} = \lambda^3 e^{-\lambda 18}$

$\ln L(\lambda) = 3 \ln(\lambda) - \lambda 18$

$$\frac{d}{d\lambda} \ln L(\lambda) = \frac{3}{\lambda} - 18 \overset{set}{=} 0$$

$$\text{or } 18\lambda = 3 \quad \text{or } \boxed{\hat{\lambda} = \frac{3}{18} \doteq 0.1667}$$

<u>Alternatively</u>, $L(\lambda) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \Sigma x_i}$

$$\ln L(\lambda) = n \ln(\lambda) - \lambda \Sigma x_i$$

$$\frac{d}{d\lambda} \ln L(\lambda) = \frac{n}{\lambda} - \Sigma x_i \overset{set}{=} 0$$

$$\text{or } n = \lambda \Sigma x_i \,, \quad \hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\bar{x}}$$

and at the data, $\hat{\lambda} = \frac{3}{6+5+7} = \boxed{\frac{3}{18} = 0.1667}$

30] The <u>parameter space</u> $\Theta = \{\underset{\sim}{\theta} \mid f(x) \text{ is a pdf or } p(x) \text{ is a pmf}\}$.

ex] If $X \sim N(\mu, \sigma^2)$ or $X \sim LN(\mu, \sigma)$

$$\Theta = \{(\mu, \sigma) \mid \mu \in \mathbb{R}, \sigma > 0\}$$

31] The maximum likelihood estimator $\underline{\hat{\theta}} = \hat{\theta}(\underset{\sim}{x}) \in \Theta$

of $\underset{\sim}{\theta}$ is the parameter value at which $L(\underset{\sim}{\theta}) = L(\underset{\sim}{\theta} \mid \underset{\sim}{x})$ attains its maximum as a function of $\underset{\sim}{\theta}$ with $\underset{\sim}{x}$ held fixed. The MLE based on the sample $\underset{\sim}{X} = (X_1, \dots, X_n)$ is $\hat{\underset{\sim}{\theta}}(\underset{\sim}{X})$.

32] If the MLE $\hat{\underset{\sim}{\theta}}$ exists, then $\hat{\underset{\sim}{\theta}} \in \Theta$. maximizing $\ln L(\underset{\sim}{\theta})$ is equivalent to maximizing $L(\underset{\sim}{\theta})$.

33] * Invariance principle:

If $\hat{\underset{\sim}{\theta}}$ is the MLE of $\underset{\sim}{\theta}$, then $\tau(\hat{\underset{\sim}{\theta}})$ is the MLE of $\tau(\underset{\sim}{\theta})$ where $\tau : \Theta \to \mathbb{R}^j$ is a function of $\underset{\sim}{\theta}$.

ex] $X_1, \dots, X_n$ iid geometric (β)

$$P(x) = \frac{\beta^x}{(1+\beta)^{x+1}}, \quad L(\beta) = \prod_{i=1}^{n} P(x_i) = \beta^{\Sigma x_i} \frac{1}{(1+\beta)^{\Sigma x_i + n}}$$

$$\ln L(\beta) = \Sigma x_i \ln(\beta) - (\Sigma x_i + n) \ln(1+\beta)$$

$$\frac{d}{d\beta} \ln L(\beta) = \Sigma x_i \frac{1}{\beta} - (\Sigma x_i + n)\frac{1}{1+\beta} \overset{set}{=} 0$$

$$\text{or } \Sigma x_i (1+\beta) = (\Sigma x_i + n)\beta$$

cancel Σx_i's from both sides

$$\text{or } \Sigma x_i = n\beta$$

$$\text{or } \hat{\beta} = \overline{X}.$$

ex) $X_1, \ldots, X_n$ iid negative binomial (r, β), r known

$$p(x) = c_x \frac{\beta^x}{(1+\beta)^{r+x}}, \quad L(\beta) = \overbrace{c}^{\prod_{i=1}^{n} c_{x_i}} \beta^{\Sigma x_i} \frac{1}{(1+\beta)^{nr + \Sigma x_i}}$$

$$\ln L(\beta) = d + \Sigma x_i \ln(\beta) - (nr + \Sigma x_i)\ln(1+\beta)$$

$$\frac{d}{d\beta} \ln L(\beta) = \Sigma x_i \frac{1}{\beta} - (nr + \Sigma x_i)\frac{1}{1+\beta} \overset{set}{=} 0$$

$$\text{or } \Sigma x_i (1+\beta) = (nr + \Sigma x_i)\beta$$

$$\text{or } \Sigma x_i = nr\beta, \quad \hat{\beta} = \frac{\overline{X}}{r}.$$

end exam 1 material

34] Often $L(\underline{\theta}) = L(\underline{\theta} \mid x_1, \ldots, x_n) = \prod_{i=1}^{n} L(\underline{\theta} \mid x_i)$.

35] For grouped data, $L(\underline{\theta} \mid x_i) =$

$F(c_j) - F(c_{j-1})$ if x_i is in interval 19.5

$(c_{j-1}, c_j] = P(X \in (c_{j-1}, c_j])$

interval	count
$(c_0, c_1]$	n_1
$(c_1, c_2]$	n_2
$\vdots$	
$(c_{m-1}, c_m]$	n_m

$$L(\underline{\theta}) = \prod_{j=1}^{m} \left[F(c_j) - F(c_{j-1})\right]^{n_j}$$

ex]

	claim size	# claims
$[0, 1000)$	Under 1000	10
	$[1000, 2000)$	5
	(2000 up)	3

Fit an $EXP(\theta)$ dist to the above table and estimate the mean.

Soln] $F(x) = 1 - e^{-x/\theta}$, $F(\infty) = 1$, $F(0) = 0$, $F(c_j) - F(c_{j-1}) = 1 - e^{-c_j/\theta} - 1 + e^{-c_{j-1}/\theta}$

$$L(\theta) = \left[F(1000) - F(0)\right]^{10} \left[F(2000) - F(1000)\right]^{5} \left[F(\infty) - F(2000)\right]^{3}$$

$$= \left[1 - e^{-\frac{1000}{\theta}}\right]^{10} \left[e^{-\frac{1000}{\theta}} - e^{-\frac{2000}{\theta}}\right]^{5} \left(e^{-\frac{2000}{\theta}}\right)^{3}$$

$$= \left(1 - e^{-\frac{1000}{\theta}}\right)^{10} \left[e^{-\frac{1000}{\theta}}\left(1 - e^{-\frac{1000}{\theta}}\right)\right]^{5} \left(e^{-\frac{1000}{\theta}}\right)^{6}$$