Lecture 2: Probability Models/Distributions

Brian Thelen
443 West Hall
bjthelen@umich.edu

Statistics 509 - Winter 2016

Lecture Overview

Brief review of relevant probability models and random variables

- Random variables
- Common distributions for random variables
  - Discrete - Bernoulli and binomial
  - Continuous - Normal, exponential, gamma, t, generalized error, and pareto/generalized pareto

Random Variables

Definition (random variable). A random variable is a function X from sample space Ω to the real numbers. We write

$$
\begin{aligned}
(X = x) &\quad \text{as shorthand for } \{w \in \Omega;\ X(w) = x\} \\
(a \le X \le b) &\quad \text{as shorthand for } \{w \in \Omega;\ a \le X(w) \le b\} \\
(X \le b) &\quad \text{as shorthand for } \{w \in \Omega;\ X(w) \le b\} \\
(X \ge a) &\quad \text{as shorthand for } \{w \in \Omega;\ X(w) \ge a\}
\end{aligned}
$$

Remark. Random variables are (typically) classified as **discrete** or **continuous**

- **Discrete** corresponds to X taking values within a discrete set (could be infinite number)
- **Continuous** corresponds to X taking values over a continuum (e.g., an interval)

Remark. The discrete vs. continuous refers to the model.

- stock/bond price
- individual/company income levels

Cumulative Distributions

Definition. The cumulative distribution function (cdf) of a discrete/continuous rv X is always the function

$$F(x) = P(X \leq x)$$

Picture: Example cdf of discrete rv/cdf of continuous rv

Remark. Note that

- cdf F is an increasing function, i.e., $F(x) \qquad F(x')$ for $x < x'$.
- As $x \downarrow -\infty$, $F(x) \rightarrow$
- As $x \uparrow \infty$, $F(x) \rightarrow$

Remark. Suppose that X is the stock price of Google at the end of this coming week. What can we say about $F(x)$ for $x < 0$.

Answer.

Quantiles

Definition. For rv X with cdf F, the quantile function is sort of a generalized inverse of F, Specifically for $0 < q < 1$,

$$F^{-1}(q) = \inf\{x : \ q \leq P(X \leq x)\}$$

or equivalently,

$$F^{-1}(q) =$$

Pictures

Moments of Distribution/Random Variable

Definition. For rv X with distribution function F, we define **central moments μ_k** as

$$\mu_k = E\left[(X - E(X))^k\right] \qquad k = 2, 3, 4, ...$$

Definition. For sample $x_1, x_2, \ldots, x_n$, define **sample mean** as

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

and we define the **central sample moments m_k** for this data as

$$m_k = \qquad k = 2, 3, 4, \ldots$$

Background: Typically assuming that sample $x_1, x_2, \ldots, x_n$ corresponds to a sampling from some process/population with an underlying distribution function F.

Summary Statistics

Background. Suppose that $X \sim F$ and have sample $x_1, x_2, \ldots, x_n$, and central moments of μ_k, m_k for $k = 2, 3, 4$.

Parameters/Statistics	Distn Parameter	Sample Statistic
Standard deviation	$\sigma = \sqrt{\mu_2}$	$\mathrm{SD}(x) =$
Skewness	$\frac{\mu_3}{(\mu_2)^{\frac{3}{2}}}$	
(Excess) Kurtosis	$\frac{\mu_4}{\mu_2^2} - 3$	

Remarks.

- Skewness is a measure of the of the distribution
- Kurtosis is a measure of how the distribution is

Discrete Random Variables

Discrete Random Variables. A discrete random variable X is totally described (with repect to probability) by defining the set of values, $\{x_i\}$, that X can take and the probability that it can take those values, i.e.,

$$p(x_i) = P(X = x_i).$$

The function p is the **probability mass function (pmf).** The **cumulative distribution function (cdf)** is given by

$$F(x) = P(X \leq x) = \sum_{i:x_i \leq x} p(x_i).$$

Expected Values/Variance/Standard Deviation

Remark. Expected value of a rv (or equivalently a probability distribution) is a

- measure of the "center" of the pmf of rv

Variance/Standard deviation of a rv is a measure of

- the spread of the pmf about the expected value

Definition. Suppose X is a discrete random variable with a pfm p. The **expected value of X** is given by

$$E(X) = \sum_x xp(x).$$

provided that $\sum_x |x|p(x) < \infty$. If sum is infinite, we say the expected value is undefined. The **variance of X** is

$$\text{Var}(X) = E\left[(X - E(X))^2\right] = E(X^2) - [E(X)]^2$$

$$E(X^2) = \sum_x x^2 p(x).$$

$$\text{SD}(X) = \sqrt{\text{Var}(X)}$$

Discrete Distribution: Bernoulli

Bernoulli Random Variables. X is a Bernoulli rv with parameter (probability) p if

- X takes only two values, 0 and 1,
- $P(X=1)=p$ and $P(X=0)=1-p$
- **Notation:** $X \sim \text{Ber}(p)$.

Definition For an event A, let $\mathbf{1}_A$ denote the indicator function, i.e.,

$$\mathbf{1}_A(w) = \begin{cases} 1 & w \in A \\ 0 & w \notin A \end{cases}$$

Remark. For an event A, $\mathbf{1}_A$ is a Bernoulli random variable with $p = P(A)$.

Expected Value/Variance of $X \sim$ Ber(p)

$$\begin{aligned} E(X) &= \\ \text{Var}(X) &= \end{aligned}$$

Discrete Distribution: Binomial

Binomial Random Variables. Suppose have independent and identically distributed (iid) Bernoulli rvs $Z_1, \ldots, Z_n$ with parameter p and $X = \sum_{i=1}^{n} Z_i$. Then

- X is **binomial** rv with parameters n and p
- **Notation:** $X \sim \text{Bin}(n, p)$
- X has pmf

$$p(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, \ldots, n$$

where (we remind reader)

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}.$$

Remark. If X is # of successes from following experiment
(1) n independent trials resulting in a success or failure
(2) probability of success for each trial is p.
Then $X \sim \text{Bin}(n, p)$.

Background on Binomial Distribution

Recall: X has pmf

$$p(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, \ldots, n$$

The justification for the above formula is:

(i) Based on the independence of the trials, the probability of any specific ordered sequence of successes/failures with exactly k successes and $(n-k)$ failures is

$$p^k (1-p)^{n-k}$$

(ii) total number of ways to distribute the k successes amongst the n trials is the same as the number of samples of size k from a population of size n (without replacement) which is $\binom{n}{k}$.

Expected Value/Variance of Binomial RVs

For $X \sim \text{Bin}(n, p)$,

$$\begin{aligned} E(X) &= \qquad\qquad\qquad\qquad . \\ \text{Var}(X) &= \end{aligned}$$

Commands in R for Binomial distribution

```
dbinom(x, size, prob)  # mass function
pbinom(q, size, prob)  # cdf
qbinom(p, size, prob)  # quantile
rbinom(n, size, prob)  # generates random deviates
```

Continuous Random Variables

- RVs are often modeled as taking values over a continuum
  - RVs are called continuous and
  - the distribution is characterized via a probability density function (pdf), $f(x)$,

$$f(x) \geq 0 \quad \text{for all x,} \quad \int_{-\infty}^{\infty} f(x) = 1$$

$$P(a < X < b) = \int_a^b f(x)\,dx$$

- The cdf of X is continuous and is given by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$$

Expected Value/Variance/St Dev of Continuous RVs

Definition. Suppose X is a continuous random variable with a pdf f. The **expected value of X** is given by

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx.$$

provided that $\int_{-\infty}^{\infty} |x| f(x)\, dx < \infty$. Otherwise we say the expectation is undefined. Provided the expected value exists, the **variance of X** is then defined as

$$\text{Var}(X) = E\left[(X - E(X))^2\right] = E(X^2) - [E(X)]^2$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx.$$

The **standard deviation of X** is

$$\text{SD}(X) = \sqrt{\text{Var}(X)}$$

Properties of Continuous RVs

Suppose X is a continous rv with pdf f and cdf F, and $a < b$ and c are real numbers. Then

(i) $P(X = c) =$

(ii)

$$\begin{aligned} P(a \leq X \leq b) &= P(a < X < b) \\ &= P(a < X \leq b) = P(a \leq X < b) \end{aligned}$$

(iii) Can write $P(a \leq X \leq b)$ in terms of cdf, i.e.,

$$P(a \leq X \leq b) = \qquad\qquad\qquad\qquad .$$

Expectation/Variance/Quantiles of Functions of RVs

Results. Suppose X is a rv and $Y = a + bX$. Then

(a) If mean $E(X)$ exists, then $E(Y) =$.

(b) If mean and variance of X exists, then
Var$(a + bX) =$.

(c) If $b > 0$, quantiles are related via $y_q =$.

(c') If $b < 0$, quantiles are related via $y_q =$.

(d) If h is a strictly increasing function and $Y = h(X)$, then quantiles of Y are given by

$$y_q =$$

(d)' If h is a strictly decreasing function and $Y = h(X)$, then quantiles of Y are given by

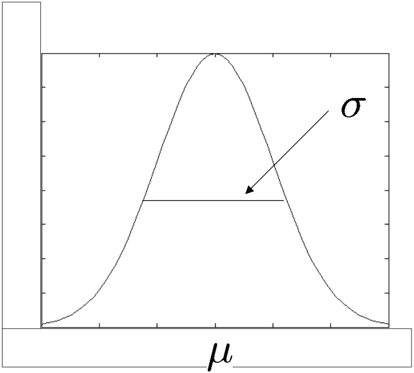
$$y_q =$$

Proof.

Continuous Distribution: Normal Distribution

Definition: Normal Distribution. A pdf f is said to correspond to a normal distribution with a mean of μ and standard deviation of σ if

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < \infty$$



Remarks/notation for normal distributions

- If X is a normal rv with parameters μ, σ^2, write $X \sim \mathcal{N}(\mu, \sigma^2)$
- $\mathcal{N}(0, 1)$ is referred to as standard normal distribution and the corresponding pdf and cdf are denoted by ϕ and Φ, i.e.,

$$\begin{aligned} \phi(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\ \Phi(x) &= \end{aligned}$$

- No nice closed form for Φ, but it is easily computed via R.

More on the Normal Distribution

Expected Value/Variance/Skewness/Kurtosis of $X \sim \mathcal{N}(\mu, \sigma^2)$

$$
\begin{aligned}
E(X) &= \mu \\
\mathsf{Var}(X) &= \sigma^2 \\
\mathsf{SD}(X) &= \sigma \\
\mathsf{Skew}(X) &= 0 \\
\mathsf{Kurt}(X) &= 0 \quad \text{excess kurtosis}
\end{aligned}
$$

Normal Distribution R-functions

```
dnorm(x, mean=0, sd=1) # density function
pnorm(q, mean=0, sd=1) # cdf
qnorm(p, mean=0, sd=1) # quantile
rnorm(n, mean=0, sd=1) # generates random deviates
```

- Can drop the "mean=" and "sd="

Plots/Histograms for Normal Distn

R-code and Output Plots

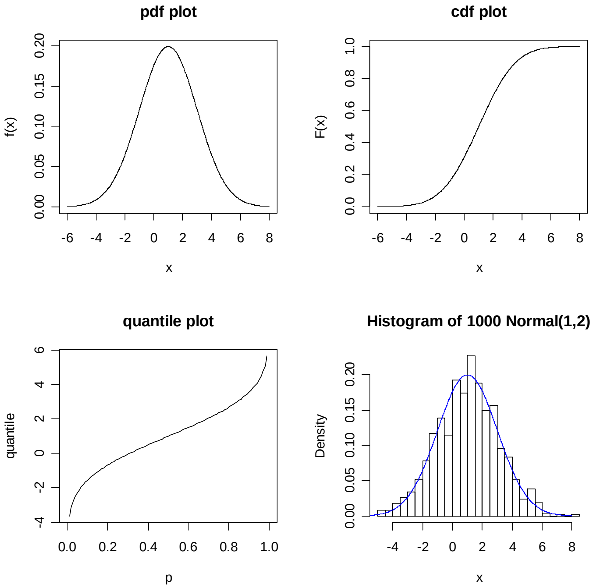
```
## Plots for Normal Case ##

mu <- 1
sigma <- 2
x <- seq(-6,8, by = .01) # vector of x-values
p <- seq(0,1, by = .01)  # vector of probabilities
n <- 1000

dnormo <- dnorm(x, mu, sigma) # pdf values
pnormo <- pnorm(x, mu, sigma) # cdf values
qnormo <- qnorm(p, mu, sigma) # quantiles
rnormo <- rnorm(n, mu, sigma) # random deviates

# Doing 3 plots and histogram
windows()
par(mfrow=c(2,2)) # setting up for a 2 x 2 arrangement of subplots
plot(x,dnormo,xlab='x',ylab='f(x)',type='l',main='pdf plot')
plot(x,pnormo,xlab='x',ylab='F(x)',type='l',main='cdf plot')
plot(p,qnormo,xlab='p',ylab='quantile',type='l',main='quantile plot')
hist(rnormo,xlab='x',breaks=25,main='Histogram of 1000 Normal(1,2)',freq=FALSE)
par(col="blue")
lines(x,dnormo)
```

Output Plots from R-code



Example.

Example Suppose $X \sim \mathcal{N}(2, 2)$ and $Y = e^X$ – find .99-quantile of Y.

Answer.

Continuous Distribution: Exponential

Definition (exponential distribution). A pdf f is said to correspond to an exponential distribution with parameter λ if it is given by

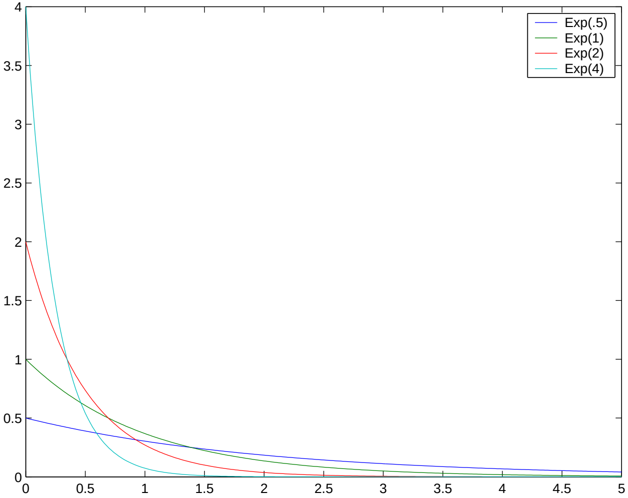
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

The cdf for $x \geq 0$ is given by

$$\begin{aligned} F(x) &= \int_0^x f(u)\,du \\ &= \int_0^x \lambda e^{-\lambda u}\,du = \qquad\qquad . \end{aligned}$$

and $F(x) = 0$ for $x < 0$. A rv X with the above pdf

- is an exponential rv, and write $X \sim \text{Exp}(\lambda)$



Plots of Exponential pdf's

More on Exponential Distribution

Expected Value/Variance of $X \sim$ Exp(λ)

$$\begin{aligned} E(X) &= \qquad\qquad\qquad\qquad . \\ \text{Var}(X) &= \\ \text{SD}(X) &= \end{aligned}$$

Exponential Distribution R-functions

- Note that rate $= \lambda$

```
dexp(x, rate=a) # density function
pexp(q, rate=a) # cdf
qexp(p, rate=a) # quantile
rexp(n, rate=a) # generates random deviates
```

Continuous Distribution: Double Exponential

Definition (double exponential distribution). A pdf f is said to correspond to a double exponential distribution with mean μ and parameter λ if it is given by

$$f(x) = \frac{\lambda}{2} e^{-\lambda|x-\mu|}$$

The cdf is given by

$$F(x) = \begin{cases} \frac{1}{2} e^{-\lambda|x-\mu|} & x < \mu \\ 1 - \frac{1}{2}\, e^{-\lambda|x-\mu|} & x \geq \mu \end{cases}$$

Remark. For rv X with above pdf, write $X \sim \mathrm{DExp}(\mu, \lambda)$.

Pictures

Expected Value/Variance of $X \sim \text{DExp}(\mu, \lambda)$

$$
\begin{aligned}
E(X) &= \qquad\qquad\qquad\qquad . \\
\text{Var}(X) &= \\
\text{SD}(X) &= \\
\text{Skew}(X) &= \\
\text{Kurt}(X) &=
\end{aligned}
$$

Double exponential Distribution R-functions

- These are in startup.R on CTools – put startup.R in R work directory
- Need to do command of `source('startup.R')`

```
ddexp(x,mu,lambda) # density function
pdexp(x,mu,lambda) # cdf
qdexp(p,mu,lambda) # quantile
rdexp(n,mu,lambda) # generates random deviates
```

Generalized Error Distributions

Definition. A rv X has a Generalized Error Distribution with parameter ν if

$$f_{ged,\nu}(x) = \kappa_\nu e^{-\frac{1}{2}\left|\frac{x}{\lambda_\nu}\right|^\nu}$$

- κ_ν and λ_ν are constants determined by ν and $\mathrm{Var}(X) = 1$ (for more details consult section 5.6 in Ruppert)
- for $\nu = 2$ have and $\nu = 1$ have
- can generalize to location-scale family, by

$$Y = \mu + \lambda X$$

  - Now have 3 parameters of μ, λ, ν
  - μ is the mean and λ^2 is the variance
  - Notation of $Y \sim \mathrm{GED}(\mu, \lambda^2, \nu)$

More on GED-Distribution

Exp Value/Variance/St Dev/Skewness/Kurtosis of GED-RVs

For $X \sim \text{GED}(\mu, \lambda^2, \nu)$

$$\begin{aligned} E(X) &= \quad . \\ \text{Var}(X) &= \\ \text{SD}(X) &= \\ \text{Skew}(X) &= \\ \text{Kurt}(X) &= \end{aligned}$$

GED-Distribution R-functions

```
dged(x, mean = 0, sd = 1, nu = 2) # density function
pged(q, mean = 0, sd = 1, nu = 2) # cdf
qged(p, mean = 0, sd = 1, nu = 2) # quantile
rged(n, mean = 0, sd = 1, nu = 2) # generates random deviat
```

- Above is from package fGarch

Continuous Distribution: Gamma

Definition (gamma distribution) A pdf f is said to correspond to a gamma distribution with parameters $\alpha, \lambda > 0$ if

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\Gamma(\alpha)$ is the gamma function

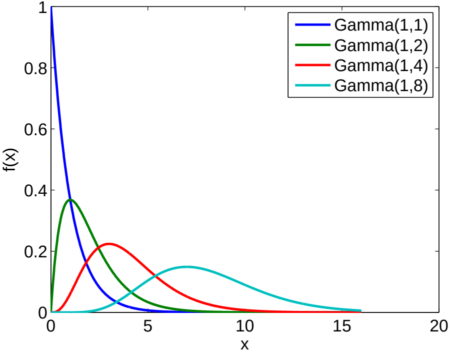
$$\Gamma(\alpha) = \int_0^\infty x^{(\alpha-1)} e^{-x}\, dx.$$

1. λ is called the scale parameter
2. α is called the shape parameter

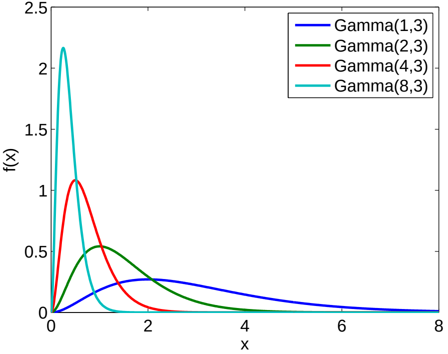
If X is a random variable with above pdf, write $X \sim \text{Gamma}(\lambda, \alpha)$.

Plots of Gamma pdf's

$\lambda = 1,\ \alpha = 1, 2, 4, 8$



$\alpha = 3,\ \lambda = 1, 2, 4, 8$



More on Gamma Distribution

Expected Value/Variance/St Dev of Gamma RVs

For $X \sim \text{Gamma}(\lambda, \alpha)$

$$\begin{aligned} E(X) &= \qquad\qquad\qquad\qquad . \\ \text{Var}(X) &= \\ \text{SD}(X) &= \end{aligned}$$

Gamma Distribution R-functions

```
dgamma(x, shape, rate=a) # density function
pgamma(q, shape, rate=a) # cdf
qgamma(p, shape, rate=a) # quantile
rgamma(n, shape, rate=a) # generates random deviates
```

- Can drop the "rate="

Properties/attributes of Gamma distribution

- **Flexible distribution – overall shape**
  - exponential is a special case with $\alpha = 1$
    - relies on $\Gamma(1) = 1$ which is easy to verify
- **Sum of iid exponential rvs is gamma**
  If $X_1, X_2, \ldots, X_n$ are iid Exp(λ), then the sum

$$Y = \sum_{i=1}^{n} X_i \sim \qquad\qquad\qquad\qquad .$$

- **Sum of iid gamma rvs is gamma**
  If $X_1, X_2, \ldots, X_n$ are iid Gamma(λ, α), then

$$Y = \sum_{i=1}^{n} X_i \sim \qquad\qquad\qquad\qquad .$$

- **Chi-square** with ν degrees of freedom is Gamma $\left(\frac{1}{2}, \frac{\nu}{2}\right)$ and corresponds to distribution of $\sum_{i=1}^{\nu} Z_i^2$ when $Z_1, Z_2, \ldots, Z_\nu$ are iid $\mathcal{N}(0, 1)$.

Continuous Distribution: t-distribution

Definition. Suppose $Z \sim \mathcal{N}(0,1)$ and $W \sim \chi^2_\nu$ are independent. Then $X = \frac{Z}{\sqrt{\frac{W}{\nu}}}$ has a t-distribution with ν degrees of freedom, and has a pdf given by

$$f_{t,\nu}(x) = \left[\frac{\Gamma\left\{\frac{\nu+1}{2}\right\}}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \right] \frac{1}{\left\{1 + \left(\frac{x^2}{\nu}\right)\right\}^{\frac{\nu+1}{2}}}$$

- General t distribution with parameter $\nu > 0$ - denote by t_ν
- Scaled t-distribution for $Y = \mu + \lambda X$, where $X \sim t_\nu$ and $\lambda > 0$. Notation is $Y \sim t_\nu(\mu, \lambda^2)$ and this is classical t
- There is a standardized t and need to be careful (see textbook), in particular $t_\nu^{std}(\mu, \sigma^2)$ corresponds to re-scaled t-distn so as to have mean 0 and variance of σ^2 for $\nu > 2$

More on t-Distribution

Exp Value/Variance/St Dev/Skewness/Kurtosis of t-RVs

For $X \sim t_\nu(\mu, \lambda^2)$

$$\begin{aligned} E(X) &= \qquad . \\ \text{Var}(X) &= \\ \text{SD}(X) &= \\ \text{Skew}(X) &= \\ \text{Kurt}(X) &= \end{aligned}$$

t-Distribution R-functions

```
dt(x, df) # density function
pt(q, df) # cdf
qt(p, df) # quantile
rt(n, df) # generates random deviates
```

- Invoke "help(TDist)" in R for more information

Heavy-Tailed Distributions

Remark. Often the rvs X will represent

- Stock price
- Bond price
- Currency exchange rate
- Insurance claims
- Aggregate price of stocks
  - SP500
  - Russell 2000
  - Dow Jones Industrial Average

Remark. From the viewpoint of risk, focus is often on the tail probabilities – for example if X is your loss (negative return), interested in

$$P(X > x) = 1 - F(x).$$

- The above is called a tail probability

Pictures

Overview of Value-at-Risk - VaR

Example. Value-at-Risk VaR The basic idea of VaR is that it helps to quantify the amount of capital needed for covering a loss in a portfolio. Consider the following:

$$\begin{aligned} P_t &= \text{value of portfolio at time } t \\ P_{t+\triangle t} &= \text{value of portfolio at time } t+\triangle t \\ R_t = \frac{P_{t+\triangle t} - P_t}{P_t} &= \text{net return at time } t+\triangle t \\ \alpha &= \text{(small) probability of not covering losses} \end{aligned}$$

The Value-at-Risk at time t, VaR_t, is defined by

$$P\left(P_{t+\triangle t} - P_t + \mathsf{VaR}_t < 0\right) = P\left(R_t + \tilde{\mathsf{VaR}}_t < 0\right) = \alpha$$

where

$$\tilde{\mathsf{VaR}}_t = \frac{\mathsf{VaR}_t}{P_t} = \text{relative Value-at-Risk}$$

More on Value-at-Risk - VaR

Remark. By previous slide,

(i) $-\text{VaR}_t$ is the α-quantile of the distn of raw return $P_{t+\triangle t} - P_t$

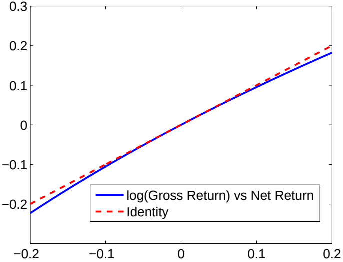
(ii) $-\tilde{\text{VaR}}_t$ is the α-quantile of the distn of return R_t

Remark. For this class we typically take logarithm of return, i.e.,

$$\tilde{R}_t \equiv \log\left[\frac{P_{t+\triangle t}}{P_t}\right] \approx R_t$$

where the approximation is justified when R_t is relatively close to 0. Note we could convert quantiles for $\tilde{R}$ to quantiles of R easily.

Log(Gross Return) vs Net Return



Formulas

Example. Suppose portfolio value is currently 100 million dollars and $\alpha = .01$.

(a) Suppose that distribution of relative return R_t is normal with a mean of .02 and a standard deviation of .03. Derive VaR and $\widetilde{\text{VaR}}$.

(b) Suppose that distribution of relative return R_t is DExp with a mean of .02 and a standard deviation of .03. Derive VaR and $\widetilde{\text{VaR}}$.

Answer.

Shortfall Distribution

- Given a level q, the CDF of the shortfall distribution is defined by

$$\Theta_q(x) = P(X \leq x | X > VaR_q)$$

- Expected shortfall

$$ES_q = E(X|X > VaR_q) = \frac{1}{q} \int_{x > VaR_q} x dF(x)$$

- Estimation of VaR_q and ES_q.

Tail probabilities

- For double-exponential distribution with mean μ and scale parameter λ, tail probability is

$$1 - F(x) = \frac{1}{2} e^{-\lambda|x-\mu|}$$

- For normal distribution, tail probability is

$$\begin{aligned} 1 - F(x) &= \int_x^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\ &\sim \frac{\sigma}{\sqrt{2\pi}x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{as } x \to \infty \end{aligned}$$

where the similar ($\sim$) notation here means that

$$\frac{1 - F(x)}{\frac{\sigma}{\sqrt{2\pi}x} e^{-\frac{(x-\mu)^2}{2\sigma^2}}} \to 1 \quad \text{as } x \to \infty.$$

Remark. In both cases, the tail probabilities go to 0 exponentially fast. There are cases where the tail probabilities go to 0 slower than exponential.

Background: Similarity of Tail Probabilities

Remarks.

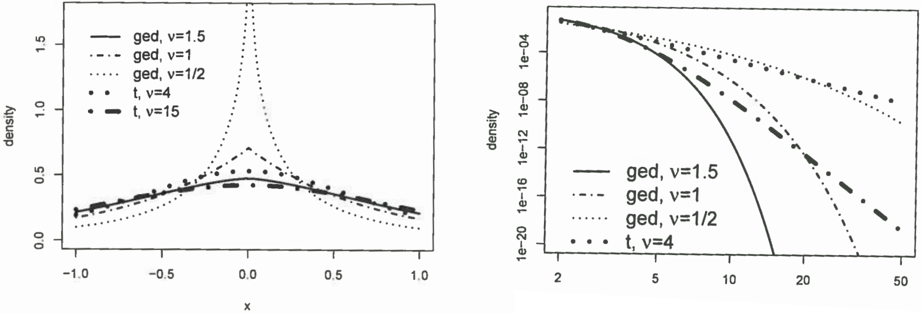
- For double exponential, normal distribution, and GED, the tail probabilities go to 0 exponentially fast.
- A number of models for financial data involve tail probabilities that go to 0 slower than exponential
  - In some widely utilized models, tail probabilities go to 0 inversely related to a polynomial – example would be

$$1 - F(x) \sim \frac{1}{x^p} \quad \text{as} \;\; x \to \infty$$

  where $p > 0$

Remark. Want to investigate for two different cdfs F and G whether tail probabilities are similar, i.e., $(1 - F(x)) \sim (1 - G(x))$. To do this we need a definition of functions being similar at ∞.

Comparison of GED with t-distribution



Implications for fitting tails vs. main body of distributions

Similarity of Tail Probabilities

Definition. Suppose have two positive functions $g(x)$ and $h(x)$ for $x \in \mathbb{R}$. We say that $g \sim h$ at infinity if there is $0 < C < \infty$ so that

$$\frac{g(x)}{h(x)} \to C \qquad \text{as } x \to \infty$$

Remark. Note that $g \sim h$ if and only if .

Example 1. Suppose $g(x)$ and $h(x)$ are polynomials, i.e.,

$$g(x) = \sum_{i=0}^{p} a_i x^i, \quad h(x) = \sum_{j=0}^{q} b_j x^j$$

where $a_p > 0, \ b_q > 0$. Then g and h have orders p and q, respectively, and $g \sim h$ if and only if

.

Similarity of Tail Probabilities - 2

Example 2. Suppose F, G are two cdfs, and $[1 - F(x)] \sim \frac{1}{x^p}$ with $p > 0$ and $[1 - G(x)] \sim Ke^{-a|x-b|^\alpha}$ with $K, a, \alpha > 0$ – then it is easy to show that

$$\lim_{x\to\infty} \frac{1 - G(x)}{1 - F(x)} =$$

Note Implication of this Example 2 result is that polynomial tail probabilities are (much) heavier than exponential tail probabilities.

Remark. Pareto distribution (and generalized Pareto distn) are distributions with “polynomial-heavy” tails.

Continuous Distribution: Pareto

Definition: Pareto Distribution. A pdf f is said to correspond to a Pareto distribution with location parameter of μ and shape parameter $a > 0$ if

$$f(x) = \frac{a\mu^a}{x^{a+1}} \qquad\qquad x > \mu$$

The cdf is given by

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(u)\, du \\ \\ &= \end{aligned}$$

Notation: If X has Pareto distribution with parameters of μ, a, write $X \sim \text{PD}(\mu, a)$.

Pictures:

Continuous Distribution: Generalized Pareto

- There is a **Generalized Pareto distribution**
- The parameterization of the generalized Pareto distribution is not a simple "generalization" of Pareto

Definition: Generalized Pareto Distribution. The pdf f for the Generalized Pareto distribution with location parameter of μ, shape parameter $\xi > 0$, and scale parameter $\sigma > 0$ is

$$f(x) = \frac{1}{\sigma}\left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{\left(-\frac{1}{\xi}-1\right)} \qquad x \geq \mu$$

The cdf is given by

$$F(x) \quad =$$

Notation: If X has generalized Pareto distribution with parameters of μ, ξ, σ, write $X \sim \text{GPD}(\mu, \xi, \sigma)$.

More on the Generalized Pareto Distn

Remark. The tail probabilities for $X \sim \text{GPD}(\mu, \xi, \sigma)$ are

$$\begin{aligned} 1 - F(x) &= \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-1/\xi} \\ &\cdot \\ &\sim \end{aligned}$$

Relationship be Pareto/Generalized Pareto It can be shown that the Pareto distribution $\text{PD}(\mu, a)$ and the generalized Pareto distn $\text{GPD}(\mu, \xi, \sigma)$ are equal when

$$\xi = \frac{1}{a} \quad \text{and} \quad \sigma = \frac{\mu}{a}$$

Generalized Pareto Distribution R-functions

- Must install fExtremes package
- Then do command of `library(fExtremes)`

```
dgpd(x, xi = 1, mu = 0, beta = 1, log = FALSE)  # density function
pgpd(q, xi = 1, mu = 0, beta = 1, lower.tail = TRUE) # cdf
qgpd(p, xi = 1, mu = 0, beta = 1, lower.tail = TRUE) # quantile
rgpd(n, xi = 1, mu = 0, beta = 1) # generates random deviates
```

- Note that "beta" is same as our "sigma"

Limit of Generalized Pareto - $\xi \to 0$

Remark. The tail distribution of GPD(μ, ξ, σ) satisfies

$$(1 - F(x)) \sim \left(\frac{\xi}{\sigma}\right)^{-\frac{1}{\xi}} \cdot \frac{1}{x^{\frac{1}{\xi}}}$$

- As $\xi \to 0$, tails becoming lighter and lighter as polynomial power $\frac{1}{\xi}$ increases
- For $\mu = 0$, the limit of the Generalized Pareto distribution as $\xi \downarrow 0$ is the exponential distribution – derived from simple limit result in mathematics that

$$\lim_{M \to \infty} \left(1 + \frac{\beta}{M}\right)^{M} = e^{\beta}$$

Remark Assuming that $\mu = 0$, can show that

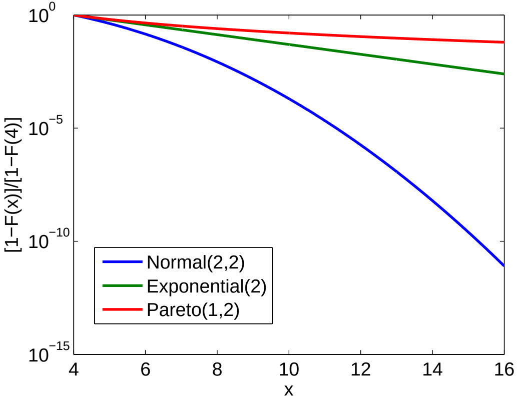
$$\lim_{\xi \to 0} (1 - F(x)) \quad =$$

i.e., tail distribution is like an .

Example: Plots of Tail Probabilities

Remark. Below is plot of tail probabilities for normal, exponential, and Pareto

- Normalized to starting point of $(1 - F(4))$, i.e., plotting conditional probability of $P(X \geq x | X \geq 4)$



Pareto Tail Distributions

Remark. Often people are most interested in the tails of a distribution (related to risk).

Definition. A rv X is said to have a Pareto (right) tail distribution if

$$1 - F(x) \sim \frac{1}{x^a}$$

Remark. Often people are most interested if distribution X looks Pareto/Generalized Pareto in the "tail" part of the distribution

- Willing to emphasize less the fit of the distribution in the middle

Example Suppose

$$F(x) = 1 - \frac{e^{\frac{1}{x^2}}}{e \cdot x^2} \qquad x > 1$$

What is the approximate tail distribution?

Answer.

Transformations of Random Variables

Linear Transformations of Normal RV

Theorem. Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$, b is any real number, $a > 0$, and $Y = aX + b$. The $Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

Corollary. Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$ and

$$Z = \frac{X - \mu}{\sigma}$$

Then Z is standard normal, i.e., $Z \sim \mathcal{N}(0, 1)$.

Linear Transformations - Uniform/Double Exp/Pareto RVs

Remark. There are similar results for uniform, double exponential, GED, t, and Pareto rvs.