Random Variables and Basic Distributional Quantities

Chapters 2 & 3

Stat 346 - Short-term Actuarial Math

Probability Review

- $\text{Pr}(E)$ will denote a number that is associated with the probability of the event E.
- Three important axioms of probability:
  1. $0 \leq \text{Pr}(E) \leq 1$
  2. $\text{Pr}(\Omega) = 1$ with Ω, the sample space.
  3. For any sequence of mutually exclusive events $E_1, E_2, \ldots$

$$\text{Pr}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} \text{Pr}(E_i), \quad \text{for } n = 1, 2, \ldots, \infty$$

- Consequence: $\text{Pr}(E^c) = 1 - \text{Pr}(E)$ for any event E.

Conditional probability

- For any two events E and A, we define

$$\Pr(E|A) = \frac{\Pr(EA)}{\Pr(A)}.$$

- Because for any event E, we have $E = EA \cup EA^c$, then

$$\begin{aligned}\Pr(E) &= \Pr(EA) + \Pr(EA^c) \\ &= \Pr(E|A) \times \Pr(A) + \Pr(E|A^c) \times \Pr(A^c).\end{aligned}$$

- Let $A_1, A_2, \ldots, A_n$ be n mutually exclusive events whose union is the sample space Ω. Then we have the Law of Total Probability:

$$\Pr(E) = \sum_{i=1}^{n} \Pr(EA_i) = \sum_{i=1}^{n} \Pr(E|A_i) \times \Pr(A_i).$$

Example: Law of Total Probability

- Consider a medical test for a particular disease in a population where:
    - 1% of people have the disease (Pr$(D) = 0.01$).
    - The test has a 99% sensitivity (Pr$(T|D) = 0.99$) and a 95% specificity (Pr$(T^c|D^c) = 0.95$).
- **Question:** What is the probability that a randomly selected individual from the population tests positive for the disease?
- Here, we will use the Law of Total Probability to find Pr(T), the probability of testing positive.

Solution: Using the Law of Total Probability

- Given:
  - $\text{Pr}(D) = 0.01$ and $\text{Pr}(D^c) = 0.99$.
  - $\text{Pr}(T|D) = 0.99$ and $\text{Pr}(T|D^c) = 0.05$.
- The Law of Total Probability states:

$$\text{Pr}(T) = \text{Pr}(T|D)\text{Pr}(D) + \text{Pr}(T|D^c)\text{Pr}(D^c)$$

- Plugging in the values:

$$\text{Pr}(T) = (0.99)(0.01) + (0.05)(0.99)$$

- Calculating the probability:

$$\text{Pr}(T) \approx 0.0594$$

- Thus, the probability that a randomly selected individual tests positive is approximately 5.94%.

Independence

- Two events E_1 and E_2 are said to be independent if

$$\Pr(E_1|E_2) = \Pr(E_1).$$

- As a consequence, we have E_1 and E_2 are two independent events if

$$\Pr(E_1 E_2) = \Pr(E_1) \times \Pr(E_2).$$

Determining Independence and Mutual Exclusivity

- Consider the following pairs of events A and B. Determine if they are independent, mutually exclusive, both, or neither.
  1. A: A six-sided die shows an even number.
     B: The same die shows a number greater than 4.
  2. A: It rains on a given day.
     B: An outdoor concert scheduled for that day is canceled.
  3. A: A randomly chosen adult male is over 6 feet tall.
     B: The same adult male has blue eyes.
  4. A: A computer part is defective.
     B: The computer part is from a batch with known issues.
  5. A: Two cards drawn from a deck are both aces.
     B: The first card drawn is an ace.
- Remember:
  - Events A and B are **independent** if $\Pr(A \cap B) = \Pr(A)\Pr(B)$.
  - Events A and B are **mutually exclusive** if $\Pr(A \cap B) = 0$; they cannot both occur at the same time.

Random variables

- A random variable will be denoted by capital letters: X.
- It is a mapping from the sample space to the set of real numbers: $X : \Omega \to \mathbb{R}$. The set of all possible values X takes is called the support of the random variable (or its distribution).
- X is a discrete random variable if it takes either a finite or at most countable number of possible values. Otherwise, it is continuous. A combination of discrete and continuous would be mixed.
- Cumulative distribution function (cdf): $F(x) = F_X(x) = \Pr(X \leq x)$.
- Survival distribution function (sdf): $S(x) = S_X(x) = \Pr(X > x)$.
- Discrete: probability mass function (pmf) $p(x) = p_X(x) = \Pr(X = x)$.
- Continuous: probability density function (pdf) $f(x) = f_X(x) = \dfrac{dF(x)}{dx}$.

Example: Exponential Distribution

- The exponential distribution is often used to model the time until an event occurs, such as the time until a machine fails.
- It is defined by its rate parameter $\lambda > 0$.
- The probability density function (pdf) is given by:

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

- The survival function, which gives the probability that the time until the event exceeds x, is:

$$S(x) = P(X > x) = 1 - F(x) = e^{-\lambda x}$$

- **Example:** Suppose the time until a light bulb fails is exponentially distributed with a mean lifetime of 1000 hours ($\lambda = \frac{1}{1000}$).
  1. Calculate the pdf for this distribution.
  2. Find the probability that a bulb lasts more than 1500 hours.

Solution: Exponential Distribution

- Given $\lambda = \frac{1}{1000}$, the pdf of the light bulb's lifetime X is:

$$f(x) = \begin{cases} \frac{1}{1000} e^{-x/1000} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

- To find the probability that a bulb lasts more than 1500 hours, use the survival function:

$$S(1500) = e^{-\frac{1500}{1000}} \approx e^{-1.5}$$

- This gives the probability that a randomly selected light bulb will last longer than 1500 hours.

Examples of random variables encountered in actuarial work

- age-at-death from birth
- time-until-death from insurance policy issue.
- the number of times an insured automobile makes a claim in a one-year period.
- the amount of the claim of an insured automobile, given a claim is made (or an accident occurs).
- the value of a specific asset of a company at some future date.
- the total amount of claims in an insurance portfolio.

Properties of distribution functions

The distribution function must satisfy a number of requirements:

- $0 \leq F(x) \leq 1$ for all x.
- $F(x)$ is non-decreasing.
- $F(x)$ is right-continous, that is, $\lim_{x \to a^+} = F(a)$.
- $F(-\infty) = 0$ and $F(+\infty) = 1$.

Some things to note:

- For continuous random variables, $F(x)$ is also left-continuous.
- For discrete random variables, $F(x)$ forms a step function.
- For a mixed random variable, $F(x)$ forms a combination with jumps and when it does, the value is assigned to the top of the jump.
- Because $S(x) = 1 - F(x)$, one should be able to deduce properties of a survival function. See page 14 of Klugman, et al.

Independent random variables

- X and Y are said to be independent if

$$\Pr(X \in C, Y \in D) = \Pr(X \in C) \times \Pr(Y \in D).$$

- Also, we have

$$\Pr(X = x, Y = y) = \Pr(X = x) \times \Pr(Y = y)$$

  and

$$f(x, y) = f_X(x) \times f_Y(y)$$

  for two independent random variables X and Y.

- Note also that if X and Y are independent and so will $G(X)$ and $H(Y)$.

Expectation

- Discrete: $\mathsf{E}[g(X)] = \sum_x g(x)p(x)$
- Continuous: $\mathsf{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$
- Linearity of expectation: $\mathsf{E}(aX + b) = a\mathsf{E}(X) + b$
- Special cases:
  - mean of X: $\mu_X = \mu = \mathsf{E}(X)$
  - variance of X: $\mathsf{Var}(X) = \sigma_X^2 = \sigma^2 = \mathsf{E}[(X - \mu_X)^2]$.
  - standard deviation of X: $\mathsf{SD}(X) = \sigma_X = \sigma = \sqrt{\mathsf{Var}(X)}$.
  - One can show that the variance can also be expressed as $\mathsf{Var}(X) = \mathsf{E}(X^2) - [\mathsf{E}(X)]^2$.
  - The k-th moment of X: $\mu'_k = \mathsf{E}(X^k)$.
  - The k-th central moment of X: $\mu_k = \mathsf{E}[(X - \mu)^k]$.
- The 2nd central moment is equal to the variance.

Other important summary quantities

- Mode of a random variable: the most likely value.
  - discrete: the value that gives the largest probability
  - continuous: the value for which the density is largest
- coefficient of variation: $\mathsf{CV}(X) = \dfrac{\mathsf{SD}(X)}{\mu_X} = \dfrac{\sigma}{\mu}$
  - dimensionless
  - useful for comparing data with different units, or with widely different means.
- skewness: $\gamma_1 = \dfrac{\mu_3}{\sigma^3}$
  - for symmetric distributions, $\gamma_1 = 0$
  - a measure of departure from symmetry
- kurtosis: $\gamma_2 = \dfrac{\mu_4}{\sigma^4}$
  - for a Normal distribution, $\gamma_2 = 3$
  - measures "flatness" or "peakedness" of the distribution

Quantiles

- quantile (sometimes called percentile) function: the inverse of the cumulative distribution function.
- The $100p$-th percentile, for $0 \leq p \leq 1$, of a distribution is any value π_p satisfying:

$$F(\pi_p-) \leq p \leq F(\pi_p).$$

  - the 50-th percentile, $\pi_{0.5}$, is the median of the distribution.
- The percentile value is not necessarily unique, unless the random variable is continuous, in which case, it is the solution to:

$$F(\pi_p) = p \quad \text{or} \quad \pi_p = F^{-1}(p).$$

- In the case it is either discrete or mixed, it is taken to be the smallest such possible value:

$$\pi_p = F^{-1}(p) = \inf\{x \in \mathbb{R} \mid F(x) \geq p\}.$$

Example: Quantiles of an Exponential Distribution

- Recall the exponential distribution with rate parameter λ, where the pdf is $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$.
- The cumulative distribution function (CDF) is $F(x) = 1 - e^{-\lambda x}$.
- The p-th quantile $(0 < p < 1)$ of a distribution is the value x_p such that $F(x_p) = p$.
- **Example:** For an exponential distribution with a mean lifetime of 1000 hours ($\lambda = \frac{1}{1000}$), find the 75th percentile.
  - This means we want to find $x_{0.75}$ such that $P(X \leq x_{0.75}) = 0.75$.

Solution: Quantiles of an Exponential Distribution

- Start with the CDF for the exponential distribution:

$$F(x) = 1 - e^{-\frac{1}{1000}x}$$

- Set $F(x_{0.75}) = 0.75$ and solve for $x_{0.75}$:

$$0.75 = 1 - e^{-\frac{1}{1000}x_{0.75}}$$

- Rearrange and solve for $x_{0.75}$:

$$e^{-\frac{1}{1000}x_{0.75}} = 0.25 \implies x_{0.75} = -1000\ln(0.25)$$

- Calculate $x_{0.75}$ to find the 75th percentile of the distribution.
- This value represents the time by which 75% of the light bulbs are expected to have failed.

Excess loss random variable

The excess loss random variable is defined to be $Y^P = X - d$, given $X > d$. Its k-th moment can be determined from

$$
\begin{aligned}
e_X^k(d) = \mathsf{E}[(X-d)^k \mid X > d] \quad &= \quad \frac{\int_d^\infty (x-d)^k f(x)dx}{1-F(d)}, \quad \text{continuous} \\
&= \quad \frac{\sum_{x_j > d}(x_j - d)^k p(x_j)}{1-F(d)}, \quad \text{discrete}
\end{aligned}
$$

When $k = 1$, the expected value

$$
e_X(d) = \mathsf{E}(X - d \mid X > d)
$$

is called the mean excess loss function. Other names used have been mean residual life function and complete expectation of life. Using integration by parts, it can be shown:

$$
e_X(d) = \frac{\int_d^\infty S(x)dx}{S(d)}
$$

Example: Excess Loss with Uniform Distribution

- Consider a policy with losses following a Uniform(0,1000) distribution.
- The policy has a deductible of 100, meaning the insurance will only pay for the part of the loss that exceeds 100.
- The excess loss random variable Y^P is defined as $Y^P = X - d$ given $X > d$, where d is the deductible.
- **Question:** What is the expected excess loss for a claim when the deductible is $d = 100$?
- We will calculate $e_X(d) = \mathsf{E}[X - d | X > d]$ for a Uniform(0,1000) distribution.

Solution: Excess Loss with Uniform Distribution

- The pdf of a Uniform(0,1000) distribution is $f(x) = \frac{1}{1000}$ for $0 \leq x \leq 1000$.
- The excess loss random variable for a deductible of 100 is $Y^P = X - 100$ given $X > 100$.
- The expected excess loss is:

$$e_X(100) = \frac{\int_{100}^{1000} (x - 100) \cdot \frac{1}{1000} dx}{1 - F(100)}$$

  where $F(100)$ is the CDF of the Uniform distribution at $x = 100$.
- Calculate $e_X(100)$ to find the expected payment by the insurance for a claim.

Exponential Distribution as a Special Case

- The exponential distribution is a special case in the context of excess loss due to its **memoryless property**.
- A random variable X is said to have the memoryless property if, for any $s, t \geq 0$, $P(X > s + t|X > s) = P(X > t)$.
- In other words, the future life expectancy of the process does not depend on how much time has already elapsed.
- For an exponential distribution with rate λ, the memoryless property simplifies the calculation of mean excess loss function.
- **Implication:** The mean excess loss function does not depend on the deductible d; it's always the mean of the original distribution.

Left censored and shifted random variable

The left censored and shifted random variable is defined to be

$$Y^L = (X-d)_+ = \begin{cases} 0, & X \le d, \\ X-d, & X > d. \end{cases}$$

Its k-th moment can be calculated from

$$\begin{aligned} \mathsf{E}[(X-d)_+^k] &= \int_d^\infty (x-d)^k f(x)dx, \quad \text{continuous} \\ &= \sum_{x_j > d} (x_j - d)^k p(x_j), \quad \text{discrete} \end{aligned}$$

Clearly, we have

$$\mathsf{E}[(X-d)_+^k] = e_X^k(d)[1-F(d)]$$

Note: for dollar events, the distinction between Y^P and Y^L is one of *per payment* versus *per loss*. See Figure 3.3 on page 26. When $k=1$, the expected value is sometimes called the stop loss premium:

$$\mathsf{E}[(X-d)_+] = \int_d^\infty S(x)dx.$$

Example: Left Censored and Shifted Random Variable

- Consider losses following a Uniform(0,1000) distribution and a deductible $d = 100$.
- The left censored and shifted random variable Y^L is defined as $Y^L = (X - d)_+$.
- Y^L represents the payment made by the insurance after the deductible is applied.
- **Question:** What is the expected payment for a claim when the deductible is $d = 100$?
- We will calculate $\mathsf{E}[(X - d)_+]$ for a Uniform(0,1000) distribution.

Solution: Left Censored and Shifted Random Variable

- The pdf of a Uniform(0,1000) distribution is $f(x) = \frac{1}{1000}$ for $0 \le x \le 1000$.
- The expected payment (after deductible) is:

$$\mathsf{E}[(X-100)_+] = \int_{100}^{1000} (x-100) \cdot \frac{1}{1000} dx$$

- Calculate $\mathsf{E}[(X-100)_+]$ to find the average payment by the insurance for a claim after applying the deductible.

Limited loss random variable

The limited loss random variable is defined to be

$$Y = X \wedge u = \begin{cases} X, & X < u, \\ u, & X \geq u. \end{cases}$$

It is sometimes called the right censored variable. Its k-th moment is

$$\begin{aligned} \mathsf{E}[(X \wedge u)^k] &= \int_{-\infty}^{u} x^k f(x)dx + u^k[1 - F(u)], \quad \text{continuous} \\ &= \sum_{x_j \leq u} x_j^k p(x_j) + u^k[1 - F(u)], \quad \text{discrete} \end{aligned}$$

With integration by parts, it can be shown that

$$\mathsf{E}[(X \wedge u)^k] = -\int_{-\infty}^{0} kx^{k-1}F(x)dx + \int_{0}^{\infty} kx^{k-1}S(x)dx.$$

Check the case when $k = 1$. Note the following important relationship:

$$X = (X - d)_+ + (X \wedge d).$$

Example: Limited Loss Random Variable

- Again consider losses following a Uniform(0,1000) distribution and an upper limit $u = 100$.
- The limited loss random variable Y is defined as $Y = X \wedge u$.
- Y represents the payment made by the insurance with the policy having an upper limit.
- **Question:** What is the expected payment for a claim when the upper limit is $u = 100$?
- We will calculate $\mathsf{E}[X \wedge 100]$ for a Uniform(0,1000) distribution.

Solution: Limited Loss Random Variable

- The pdf of a Uniform(0,1000) distribution is $f(x) = \frac{1}{1000}$ for $0 \leq x \leq 1000$.
- The expected payment (with an upper limit) is:

$$\mathsf{E}[X \wedge 100] = \int_0^{100} x \cdot \frac{1}{1000} dx + 100 \cdot \left(1 - \frac{100}{1000}\right)$$

- Calculate $\mathsf{E}[X \wedge 100]$ to find the average payment by the insurance for a claim considering the upper limit.

Risk measures

- A risk measure is a mapping from the random variable representing loss (risk) to the set of real numbers.
- It gives a single value that is intended to provide a magnitude of the level of risk exposure.
- Notation: $\rho(X)$
- Properties of a coherent risk measure:
  - subadditive: $\rho(X+Y) \leq \rho(X) + \rho(Y)$.
  - monotonic: $\rho(X) \leq \rho(Y)$ if $X \leq Y$ for all posssible outcomes.
  - positive homogeneous: $\rho(cX) = c\rho(X)$ for any positive constant c.
  - translation invariant: $\rho(X+c) = \rho(X) + c$ for any positive constant c.

Possible uses of risk measures

- Premium calculations
- Determination of risk/economic/regulatory capital requirements
- Reserve calculations
- Internal risk management
- Financial reporting e.g. meeting regulatory requirements for financial reporting (Basel Accord II)

Value-at-Risk measure

The value-at-risk of X at the $100p\%$ level, denoted by $\mathsf{VaR}_p(X) = \pi_p$, is the $100p$-th percentile (or quantile) of the distribution of X and is the solution to

$$\mathsf{Pr}(X > \mathsf{VaR}_p(X)) = 1 - p.$$

Some examples:

- Normal distribution: $\mathsf{VaR}_p(X) = \mu + \sigma\Phi^{-1}(p)$
- Lognormal distribution: $\mathsf{VaR}_p(X) = \exp[\mu + \sigma\Phi^{-1}(p)]$
- Exponential distribution: $\mathsf{VaR}_p(X) = -\theta \log(1 - p)$
- Pareto distribution: $\mathsf{VaR}_p(X) = \theta[(1 - p)^{-1/\alpha} - 1]$

Some remarks:

- Value-at-risk is monotone, positive homogeneous and translation invariant, but not necessarily subadditive. See Example 3.13.

Tail Value-at-Risk measure

The tail value-at-risk of X at the $100p\%$ security level is defined to be

$$\mathsf{TVaR}_p(X) = \mathsf{E}(X \mid X > \mathsf{VaR}_p(X)) = \frac{\int_{\pi_p}^{\infty} x f(x) dx}{1 - F(\pi_p)} = \frac{\int_{\pi_p}^{\infty} x f(x) dx}{1 - p}.$$

It is the expected value of the loss, conditional on the loss exceeding the quantile (or VaR). Other formulas for TVaR:

- $\mathsf{TVaR}_p(X) = \frac{\int_p^1 \mathsf{VaR}_u(X) du}{1-p}$
- $\mathsf{TVaR}_p(X) = \mathsf{VaR}_p(X) + e(\pi_p)$
- $\mathsf{TVaR}_p(X) = \pi_p + \frac{\mathsf{E}(X) - \mathsf{E}(X \wedge \pi_p)}{1-p}$

Some remarks:

- Other terms used are: conditional tail expectation (CTE), tail conditional expectation (TCE), expected shortfall (ES).
- The tail value-at-risk measure is considered coherent.

Tail value-at-risk for some distributions

Some examples:

- Normal: $\text{TVaR}_p(X) = \mu + \sigma \dfrac{\phi[\Phi^{-1}(p)]}{1-p}$
- Lognormal: $\text{TVaR}_p(X) = \exp(\mu + \sigma^2/2) \dfrac{\Phi(\sigma - \Phi^{-1}(p))}{1-p}$
- Exponential: $\text{TVaR}_p(X) = \theta - \theta \log(1-p) = \theta + \text{VaR}_p(X)$
- Pareto: $\text{TVaR}_p(X) = \text{VaR}_p(X) + \dfrac{\text{VaR}_p(X) + \theta}{\alpha - 1}$

Examples

Spend time in class with help from neighbors. Find the $\text{VaR}_{.95}(X)$ and $\text{TVaR}_{.95}(X)$ for each of the following distributions:

- Normal Distribution with Mean $(\mu) = 100$, Standard Deviation $(\sigma) = 15$.
- Uniform Distribution with Lower bound $(a) = 50$, Upper bound $(b) = 200$.
- Exponential Distribution with Rate $(\lambda) = 0.03$.

Beware of risk measures

- Note that a risk measure provides a single value to quantify the level of risk exposure. There is possibly no single risk measure that can provide the whole picture of the danger of the exposure.
- It is therefore important to be cautious of the uses and limitations of the risk measure being used: make sure you understand what the risk measure quantifies which aspect of the risk.
- Some risk measures are based also on the model being used; beware of model risk together with the uncertainty of the parameter used in the model.
- Downfall of hedge fund LTCM in 1998: had a sophisticated VaR-based risk management system in place but errors in parameter estimation, unexpected large market moves, vanishing liquidity contributed to downfall.