Expertise. Insight. Solutions.

Exam Content Outline

Property and Casualty Predictive Analytics (PCPA)

Overview

The CAS Property and Casualty Predictive Analytics (PCPA) represents a combination of opportunities for candidates to demonstrate their knowledge and skills in this area. These requirements will be part of the ACAS pathway and have candidates demonstrate a foundational level of expertise in predictive analytics through:

- Passing a computer-scored exam designed to measure a sample of knowledge and skills.
- Passing a predictive analytics project that includes data exploration, model construction and refinement, interpretation of results, and submission of a technical summary of these project activities in response to a business question.

The goal of the content outline is to provide an understanding of the candidate expectations and how their performance will be evaluated.

CAS recognizes that modern actuaries use a variety of software and programming languages to conduct analyses including predictive analytics. The references included in this list were selected because the content is relevant to the knowledge and skills assessed on the PCPA exam, not to communicate a preferred or recommended programming language that may be used in the examples. The CAS has developed the PCPA exam and project to be applicable to predictive analytics practice in property and casualty insurance. Candidates should review the domains and tasks in the Content Outline to best understand what is covered in the PCPA.

Exam Item Types

Candidates may see the following item types in the CAS examinations. Candidates should become familiar with these item types, and sample questions will be available in 2024.

Multiple Choice

Multiple answer choices are presented after a problem with only one correct answer.

Multiple Selection

Multiple answer choices are presented after a problem with more than one correct answer.

Matching

Content columns presented after a problem where the candidate must correctly match content from one column to another.

Fill in the Blank

A blank section is presented after a problem where the candidate must input the correct value.

Expertise. Insight. Solutions.

Project Performance Tasks

Candidates will be provided with a business problem and data to respond to a series of tasks that include identifying, building, and improving a GLM model, interpreting diagnostics, recommending a model, and communicating characteristics of the model to technical and non-technical audiences. The deliverable for the PCPA project will include submitting technical characteristics from the project, a summary report, and responses to a series of questions about the documents upon submission.

Cognitive Level

Each task is tied to a certain type of mental operation or thinking skill, which is called the cognitive level. ACAS and FCAS use four cognitive levels, and every examination item is authored to address both the task and one of the following cognitive levels paired with that task.

Remember: 5-10%

Tests the ability of the candidate to recall or remember knowledge or facts.

Understand and Apply: 55-60%

Measures the candidate's ability to understand and apply ideas and concepts to new situations.

Analyze and Evaluate: 35-40%

Requires the candidate to analyze information, combine concepts/ideas, and justify a position resulting from that combination.

Create: 0-5%

Requires the candidate to synthesize conclusions by evaluating the validity of ideas and concepts.

Expertise. Insight. Solutions.

Exam Format Details

The PCPA exam is a two-hour exam that consists of 40 multiple-choice questions and is administered as a computer-based test (CBT). For additional details, please refer to the Syllabus of Basic Education's registration and administration information.

Candidates take this exam after Exams MAS-I, MAS-II, and 5 within the ACAS credentialing pathway, and therefore it is assumed that they have foundational knowledge of predictive analytics concepts and applications. This exam is a prerequisite for project component, which is described in the next section of this document.

Each response will be scored as correct or incorrect and candidate performance will be judged by comparing their overall total score (i.e., number of correct responses) against a passing standard set by actuarial subject matter experts.

As part of the computer-based testing process, a few pilot questions will be randomly placed in the exam. These pilot questions are included to judge their effectiveness for future exams, but they will NOT be used in the scoring of this exam. All other questions will be considered in the scoring. All unanswered questions are scored incorrect. Therefore, candidates should answer every question on the exam. Candidates are advised to answer each question to the best of their ability, independent from how they have answered other questions on the exam.

Because the exam will be offered continuously, each candidate will receive a test form composed of questions selected from a pool of questions. Statistical scaling methods are used to ensure within reasonable and practical limits that, during the same testing period of a few days, all forms of the test are comparable in content and passing criteria. The methodology that has been adopted is used by many credentialing programs that give multiple examination forms. Preliminary results will be released immediately following administration and confirmed following each testing window.

Expertise. Insight. Solutions.

Project Format Details:

The project is an offline, independent activity that is part of the ACAS credentialing pathway. Candidates will be eligible to participate in the project once they pass the PCPA exam in which they demonstrate a foundational knowledge and skills with respect to predictive analytics.

The project assumes knowledge of predictive analytic statistical methods as covered in Exams MAS –I, MAS –II, and 5 as well as the necessary statistical software and coding skills to perform modeling tasks.

Candidates will be able to complete the project in one of four windows throughout the year:

PROJECT START	SUBMISSION DEADLINE	GRADE RELEASE DATE
March 16	March 31	May 31
June 15	June 30	August 31
September 15	September 30	November 30
December 16	December 31	February 28/29

At the beginning of the project window, candidates will receive the following information about their task:

- Statement of the business problem,
- One or two datasets to explore and use in model construction,
- Scope parameters for the project, and
- Guidelines as to what should be submitted for grading.

Candidates will use this information to create a predictive analytics model to address the business problem. They will be responsible for submitting a brief technical report (max of 750 words) that describes how they:

- Conducted an exploratory data analysis, identified and managed any issues within the data (e.g., variables that needed to be transformed, outliers, data errors, or missing data), and identified appropriate target and predictor variables.
- Created a GLM (i.e., Gamma, link, polynomial, Poisson, binomial, log-normal, Tweedie) to address the business problem, evaluated model performance, and used the evaluation results to improve model performance.
- Interpreted the findings of the predictive analytics model from a technical perspective but also for the business decision to be made.

In addition to this technical summary, candidates will submit the predictive analytics code they created in R, Python, or SAS, up to 1 table or 2 supporting graphics as appendices to the report, and an attestation that they completed the project work independently. Although this will not be scored, it may be reviewed for security purposes to ensure it is original work. In addition, at the point of submission, candidates will be asked to respond to questions about the results of their analysis and information in the summary report.

Because the project will be offered as an offline task and across multiple windows throughout the year, each candidate will receive a specific project selected from a pool of projects. Each project will be constructed in accordance with the same framework and be judged to be of comparable difficulty with respect to the project scope and tasks.

Expertise. Insight. Solutions.

Domain Details

The table below shows the skills candidates will demonstrate via the exam and project. Each domain is listed with a weight (percentage) that indicates how the candidate's performance in that area will be weighted on the exam (percent of items targeting the section) and project (via the scoring rubric).

DOMAIN	EXAM	PROJECT
Dealing With Data	35%	30%
Model Diagnostics & Selection	35%	30%
Model Interpretation & Presentation	30%	40%

A. Dealing with Data

Candidates should be able to evaluate the dataset and manipulate it so it can be used in a predictive analytics model.

TASKS
1. Gather and assess the relevance of information from stakeholders in actuarial analysis.
2. Import, manipulate, and evaluate datasets using generally available programming languages and software packages (e.g., .csv file).
3. Evaluate the need for variable transformation and apply appropriate transformations to data.
4. Identify and appropriately manage outliers and missing data.
Readings: - Gellman and Unwin - Grolemund and Wickham - McKinney

Expertise. Insight. Solutions.

B. Model Diagnostics & Selection

Candidates should be able to create and refine a GLM model.

TASKS
1. Create and run a GLM model.
2. Evaluate and improve a GLM model (e.g., create and interpret diagnostics, conduct cross-validation, incorporate offsets, mitigate multicollinearity issues, avoid under and overfitting) for the data provided and business goals.
Readings: - Charpentier - De Jong and Heller - Fannin - Frees, Derrig and Meyers - Venables and Ripley

C. Model Interpretation & Presentation

Candidates should be able to interpret the findings from a predictive analytics model and present their findings to technical and non-technical audiences.

TASKS
1. Create and interpret statistical/tabular and graphical/visual representations of data.
2. Communicate project technical information, including details on methodologies, modeling decisions, and interpretation of output.
3. Communicate project findings to non-technical audiences, including the implications on business outcomes or decisions.
Readings: - Cairo - Few - Knaflic

Expertise. Insight. Solutions.

D. Success Criteria for the Project

PERFORMANCE EVALUATION CRITERION	DOMAIN (TASK)
Candidate describes how any variables were transformed and why.	Dealing with Data (A-3)
Candidate describes any anomalous characteristics of the data (e.g., outliers, missing data) and how they were addressed (e.g., trimming, imputation).	Dealing with Data (A-4)
Candidate's model holds up to an out-of-sample model performance check.	Model Diagnostics & Selection (B-1)
Candidate correctly interprets diagnostic characteristics (e.g., AIC, BIC), spurious relationships, multicollinearity, correlated variables, and demonstrates that they understand how to use model diagnostics to improve model fit and performance.	Model Diagnostics & Selection (B-2)
Candidate describes which model they selected and why.	Model Diagnostics & Selection (B-2)
Candidate generates and provides technical output.	Model Diagnostics & Selection (B-2)
Candidate explains what they did to iteratively check and build their model, including which diagnostics (e.g., coefficient in/out) they used and why.	Model Diagnostics & Selection (B-2)
Candidate provides and justifies their choice of data presentation and related visual presentation, including such things as: proper labeling; clarity of purpose; appropriate for the data, model, audience, and business question; text and visualization(s) support each other.	Model Interpretation & Presentation (C-1)
Candidate describes what method(s) they used and why.	Model Interpretation & Presentation (C.2)
Candidate describes why a variable was or was not included in the model.	Model Interpretation & Presentation (C.2)
Candidate's model should work and be appropriate to effectively address the business question.	Model Interpretation & Presentation (C.3)
Candidate makes an effective persuasive argument to a non-technical audience regarding how their model addresses the business question.	Model Interpretation & Presentation (C.3)

Expertise. Insight. Solutions.

Complete Text References for PCPA

CAS recognizes that modern actuaries use a variety of software and programming languages to conduct analyses including predictive analytics. The references included in this list were selected because the content is relevant to the knowledge and skills assessed on the PCPA exam, not to communicate a preferred or recommended programming language that may be used in the examples. The CAS has developed the PCPA exam and project to be applicable to predictive analytics practice in property and casualty insurance. Candidates should review the domains and tasks in the Content Outline to best understand what is covered in the PCPA.

Text references are alphabetized by the citation column.

Citation	Abbreviation	Domains/ Tasks	Source
Actuarial Standards Board of the American Academy of Actuaries, "Actuarial Standard of Practice No. 23, Data Quality," December 2016.	ASOP 23	A	**OP**
Actuarial Standards Board of the American Academy of Actuaries, "Actuarial Standard of Practice No. 56, Modeling," December 2019.	ASOP 56	A	**OP**
Cairo, A., *How Charts Lie: Getting Smarter about Visual Information*, Norton, 2020.	Cairo	C	**B**
Charpentier, A., *Computational Actuarial Science with R, 1st Edition*, Chapman and Hall/CRC, 2015. Pages 165-183	Charpentier	B	**B**
De Jong, P., and Heller, G., *Generalized Linear Models for Insurance Data*, Cambridge, 2008. (Chapters 5, 6, 8 and related code in Appendix)	De Jong and Heller	B	**B**
Fannin, B., *R for Actuaries and Data Scientists with Applications to Insurance*, ACTEX Learning, 2020. Chapter 6: Basic Visualization Chapter 15: GLMs Chapter 19: Model Selection Chapter 18: Data pre-processing	Fannin	B	**B**
Few, S., *Show Me the Numbers: Designing Tables and Graphs to Enlighten, 2nd Edition, Analytics Press, 2012.*	Few	C	**B**
Frees, E. W.; Derrig, R. A.; & Meyers, G., *Predictive Modeling Applications in Actuarial Science: Volume 1, Predictive Modeling Techniques (International Series on Actuarial Science)*, Cambridge University Press, 2014. Chapter 6: Frequency and Severity Models	Frees, Derrig and Meyers	B	**B**

Gelman, A. and Unwin, A., "Infovis and Statistical Graphics: Different Goals, Different Views," 2012.	Gellman and Unwin	A	**OP**
Grolemund, G, and Wickham, H., *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*, 1st Edition, 2017.	Grolemund and Wickham	A	**OP**
Knaflic, C., *Storytelling with Data: A Data Visualization Guide for Business Professionals, 1st Edition*, John Wiley & Sons, Inc., 2015.	Knaflic	C	**B**
McKinney, W, *Python for Data Analysis, 3rd Edition,* O'Reilly Media, Inc., 2022	McKinney	A	**OP**
Venables, W., and Ripley, B., .(2002). *Modern Applied Statistics with S., 4th Edition*, Springer, 2002., pp. 172-176	Venables and Ripley	B	**OP**

Source Key

B	Book – may be purchased from the publisher or bookstore.
BO	Book (Optional) – may be purchased from the publisher or bookstore.
DSK	Material included in the Digital Study Kit.
NEW	Indicates new or updated material.
OP	All text references marked as Online Publications will be available on a web page titled Complete Online Text References.
SK	Material included in the Study Kit.

Expertise. Insight. Solutions.

Examination Discipline and Security Policies specific to the Project

By participating in the project, candidates agree that project activities will be done independently with no hands-on assistance. The use of online resources, including but not limited to Google, Stack Overflow, Reddit, r-project, and Python.org, is permissible. Candidates may also consult with colleagues for general questions and advice; however, project instructions and datasets are proprietary to The CAS Institute and should not be shared. While Excel can be used to perform rudimentary tasks involving the data, we expect candidates to use Python, R, or SAS to do exploratory data analysis, data wrangling, dataset merging, model development, and model performance diagnostics. All code, documentation, exhibits, and the final report must be produced by the candidate. A candidate's submitted code should be able to reproduce their data scrubbing and modeling results programmatically.

The CAS Examination Discipline Policy also applies to the project with the full version available online at 2023Syllabus.pdf (casact.org) (see pp. 29-30). Candidates are advised to review the policies and potential disciplinary actions for violating these policies prior to registering to complete the project. A section of the Discipline Policy is reproduced here:

> ***"Candidates must not give or receive assistance of any kind during or after the examination. Any cheating, attempt to cheat, assisting others to cheat, participating therein, or engaging in improper conduct such as noted in the CAS Examination Discipline Policy is a serious violation and will result in serious consequences. At the sole authority of the Vice President-Admissions these consequences may include disqualification of the candidate's exam, and a ban from sitting for any CAS Examinations for up to two calendar-years after the improper conduct occurred. In addition, the Vice President-Admissions may refer the matter to the CAS Discipline Committee to consider even more severe consequences, including a permanent ban from sitting for CAS examinations." (CAS Examination Discipline Policy, p. 29)***

Examination security practices specific to the Project

Because the project is completed outside a secure examination environment, the CAS will use additional approaches to evaluate whether candidates are completing the project independently as well as the outputs of the work. These security practices may include, but are not limited to:

- Candidates completing a "Candidate Agreement" form when submitting their project materials where the candidate attests to the independence of their work and acknowledges the potential consequences if it is determined that they did not complete their work independently.
- Candidates responding to a set of questions about the project data, processes, and/or outputs at the point of submission.
- Graders running the code submitted with the project to evaluate whether it produces the outputs included in the summary report.
- Graders using plagiarism software to evaluate summary reports.